AFATL-TR-90-63

(2)

The Design of Phase-only Filters for Optical Correlations

Robert R Kallman

UNIVERSITY OF NORTH TEXAS
POST OFFICE BOX 5116
DENTON, TEXAS 76203-5116

JULY 1990

FINAL REPORT FOR PERIOD JUNE 1986 - DECEMBER 1989

AIR FORCE ARMAMENT LABORATORY
Air Force Systems Command ▮ United States Air Force ▮ Eglin Air Force Base, Florida

NOTICE

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the United States Government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The AFATL STINFO program manager has reviewed this report, and it is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

Thomas R. Callen

THOMAS R. CALLEN, Lt Col, USAF
Chief, Advanced Guidance Division

If your address has changed, if you wish to be removed from our mailing list, or if the addressee is no longer employed by your organization, please notify AFATL/AGA, Eglin AFB FL 32542-5434.

Copies of this report should not be returned unless return is required by security considerations, contractual obligations, or notice on a specific document.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
		FINAL 1 Jun 86 to 31 Dec 89

4. TITLE AND SUBTITLE
The Design of Phase-only Filters for Optical Correlators

5. FUNDING NUMBERS
C F08635-86-K-0139
PR 2305E205

6. AUTHOR(S)
Robert R. Kallman
The program was managed for the Armament Laboratory by Dr Dennis H. Goldstein

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Department of Mathematics
University of North Texas
P.O. Box 5116
Denton, Texas 76203-5116

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Advanced Guidance Division
Air Force Armament Laboratory
Eglin AFB, Florida 32542-5434

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFATL-TR-90-63

11. SUPPLEMENTARY NOTES
Availability of report is specified on verso of front cover.

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release;
distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

In 1984-1985 the author formulated the design of phase-only filters for optical correlators as a complicated nonlinear nondifferentiable optimization question in a very large number of variables. The first objective of this program was to find effective mathematical tools for the solution of this class of optimization problems and to implement them in computer codes. The phase-only filters designed by the techniques developed during the course of this program can be packed with a great deal of information, give good target recognition in the presence of large amounts of additive zero mean Gaussian noise, and objectively perform up to two orders of magnitude better in target recognition and discrimination than filters designed by other techniques. These desirable characteristics continue to hold even after the phase-only filters have been optimally discretized for implementation in actual devices. The second goal of the program was to use the mathematical tools and computer codes developed herein for phase-only filter design when a warm target resides in a warm background and the average temperature of the target and the average temperature of the background are close.

14. SUBJECT TERMS
Optical Spatial Filters
Optical Pattern Recognition
Optical Correlation

15. NUMBER OF PAGES
115

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	SAR

PREFACE

This program was conducted by Robert R. Kallman, Department of Mathematics, University of North Texas, P.O. Box 5116, Denton, Texas 76203-5116 under contract F08635-86-K-0139 with the Air Force Armament Laboratory, Eglin Air Force Base, Florida 32542-5434. Dennis Goldstein, AFATL/AGA, managed the program for the Air Force Armament Laboratory. The work was conducted during the period from June 1, 1986, through December 31, 1989.

The author wishes to give special thanks to Dennis Goldstein for his patient interest and support during the course of this work and for many informative discussions on electro-optics and target recognition. The author also wishes to thank Rick Wehling, AFATL/SAI, for partial support during the last fifteen months of this project. A conversation with David Flannery was most helpful in clarifying the ideas of Section IX.

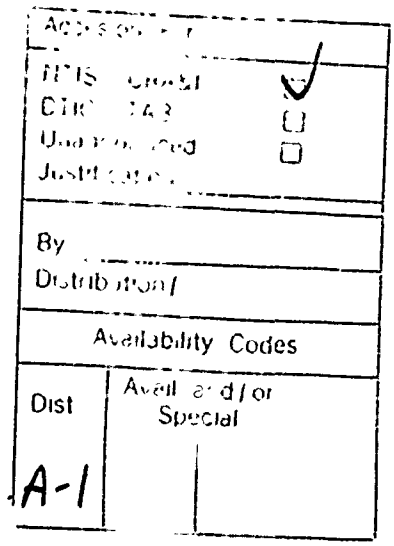


TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FIGURES (CONTINUED)

LIST OF FIGURES (CONCLUDED)

LIST OF TABLES

LIST OF TABLES (CONCLUDED)

SECTION I

INTRODUCTION

Addressable optical correlators might well find uses in target recognition and discrimination. The work reported on is concerned with some of the hard technical details involved in phase-only filter design for addressable optical correlators. The work had two major objectives. In the 1984 – 1985 period the author formulated the design of phase-only filters for optical correlators as a complicated nonlinear nondifferentiable optimization question in a very large number of variables. The latest edition of such a formulation is presented in some detail in Section II. In the same period the author observed a heretofore unnoticed, or at least apparently unpublished, difficulty in the optical correlation process using phase-only filters. This difficulty arises if a warm target resides in a warm background and the average temperature of the target and the average temperature of the background are close. This difficulty is spelled out in Section IV. A preliminary but necessary detailed technical discussion of the correct computer simulation of the optical correlation process is given in Section III. A description of the target imagery used in this study is also given in Section III. Section V gives a much more detailed discussion of the goals of this work as well as a general description of the mathematical solution of the complicated optimization problem involved in phase-only filter design. Some of the mathematical details of this approach are given in The Appendix. Four solutions to the aforementioned difficulty in the use of phase-only filters for optical correlation are given in Section VI, Section VII, Section VIII, and Section IX. What currently appears to be the favored solution can be found in Section VI, though the very recent solution given in Section VII

holds great promise. The idea propounded in Section VIII gives a theoretically close to ideal solution to the phase-only filter problem but requires a modification to the optical correlator architecture. The idea presented in Section IX requires no modification in the correlator architecture, but instead requires a change in the algorithm. It also suffers from the major difficulty of not being translation invariant (the others are, at least to a good approximation), but might prove useful for pointing and tracking. Conclusions are given in Section X.

SECTION II

THE FORMULATION OF PHASE-ONLY FILTER DESIGN AS A MATHEMATICAL OPTIMIZATION QUESTION

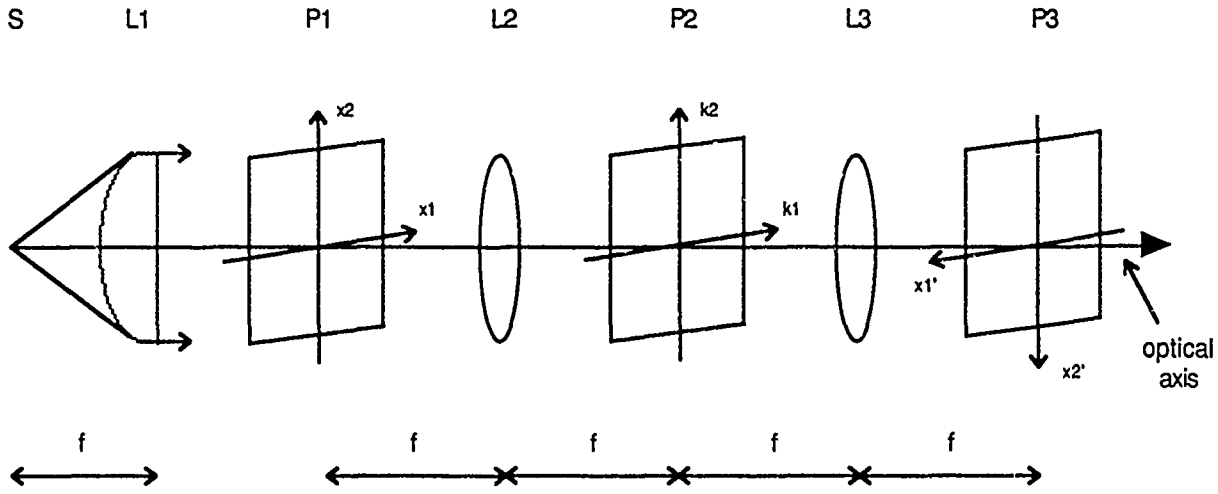
The standard correlation setup in the plane R^2 using discriminant functions will be sketched. References 1 – 7 contain exhaustive treatments of the background on optical correlation necessary to understand this section. For any pair of complex-valued square-integrable functions f and g defined on the plane, define the symbol $\langle f,g \rangle$ to be the usual complex inner product $\int_{R^2} f(y)g^*(y)dy$. Here $y = (y_1,y_2)$ is a generic point in the plane, and g^* is the complex conjugate of g. The symbol $\langle f,g \rangle$ is linear in f, conjugate linear in g, and $\langle f,g \rangle = \langle g,f \rangle^*$. Note that $\langle f,f \rangle$ is the square of the usual L^2-norm of f. For any pair of points $x = (x_1,x_2)$ and $y = (y_1,y_2)$ in the plane and any function g on the plane, let $g_x(y) = g(y - x)$. If g is initially viewed as being centered over the origin (0,0), then g_x should be viewed as being g translated so as to be centered over x. Any complex-valued square-integrable function f on the plane may be considered to be an image. This level of generality will prove useful in later sections. A good discriminant function h is a complex-valued square-integrable function on the plane which has the property that the magnitude of $\langle f,h_x \rangle$ (or equivalently $|\langle f,h_x \rangle|^2$) is relatively large if there is an object of interest in f near x and has the property that the magnitude of $\langle f,h_x \rangle$ (or equivalently $|\langle f,h_x \rangle|^2$) is relatively small otherwise. Let $\mathcal{F}$ denote the two-dimensional Fourier transform operator. It is a simple consequence of the Parseval theorem that $\langle f,h_x \rangle = \mathcal{F}^{-1}(\mathcal{F}(f)\cdot\mathcal{F}(h)^*)(x)$. This formula indicates the manner in which a

discriminant function can be implemented in an optical correlator of standard design. This implementation is summarized in the following algorithm:

(i) input the coherent image f;
(ii) use a lens to compute the Fourier transform $\mathcal{F}(f)$;
(iii) multiply $\mathcal{F}(f)$ (using a transparency, spatial light modulator, etc.) by the filter $\mathcal{F}(h)^*$ corresponding to the discriminant function h;
(iv) use a lens to compute the inverse Fourier transform $\mathcal{F}^{-1}(\mathcal{F}(f)\cdot\mathcal{F}(h)^*)$ of this product;
(v) use a detector to measure the light intensities $|\langle f,h_x\rangle|^2$.

Figure 1 contains a schematic of an optical correlator of standard design. A particular choice of device used to carry out the crucial multiplication of $\mathcal{F}(f)$ by the filter $\mathcal{F}(h)^*$ in step (iii) imposes severe constraints on admissible choices of the filter $\mathcal{F}(h)^*$. These constraints must be taken into account in filter design.

In a properly designed optical correlator the pixels on the detector should be in a one–to–one correspondence (up to scaling) with the same set of pixels in the input plane. Input devices and detectors are, of course, bounded in extent. Therefore all images encountered in practice of necessity must vanish outside the bounded region in the plane corresponding to the face of the input device. This basic fact must be taken into consideration in filter design. It may be assumed that the pixels on the detector correspond to a subset, perhaps all, of the pixels on the input device. There is no point in making the face of the detector any larger. Having chosen a carefully crafted filter $\mathcal{F}(h)^*$, there then will be something of interest in the image f at



S = Source
L1 = Collimator Lens
P1 = Input Plane
L2 = Fourier Transforming Lens
P2 = Fourier (Filter) Plane
L3 = Fourier Transforming Lens
P3 = Correlation (Output) Plane

Figure 1. A Schematic of an Optical Correlator of Standard Design

those places x where $|\langle f,h_x\rangle|^2$ is relatively large and nothing of interest at those places x where $|\langle f,h_x\rangle|^2$ is relatively small.

Next, a mathematical formulation of the intrinsic performance or signal-to-noise ratio of a discriminant function h, implemented in an optical correlator of standard design, against a training set of images will be given. This mathematical formulation will be nothing but a numerical measure of the most straightforward and naive properties that a good discriminant function should possess. Let $f_1,\ldots,f_n,\ldots,f_m$, the training set, be a finite sequence of images. The images can be ordered so that $f_1,\ldots,f_n$ will be true targets, usually centered over the origin (0,0), and $f_{n+1},\ldots,f_m$ will be false targets which must not be confused with $f_1,\ldots,f_n$. In this formulation there must always be at least one true target, but there need not be any false targets, in which case n will equal m. For each integer i between 1 and n, let B_i be a box or any other region in the correlation plane, usually but not necessarily close to the origin (0,0), but always containing the origin (0,0). In most instances each B_i $(1 \leq i \leq n)$ certainly should be a subset of the detector face close to the origin (0,0). In order for h to be a good discriminant function the largest of the measured optical intensities $|\langle f_i,h_x\rangle|^2$ (x in B_i) in the correlation plane should be relatively large for each integer i between 1 and n. Here the best choice of x in B_i may well vary with the choice of the integer i between 1 and n. In addition, the smallest of these relatively large signals should be as large as possible. In addition, let B_i be empty $(n + 1 \leq i \leq m)$ for convenience. Finally, let R_i be a region in the correlation plane which at least contains the detector face and B_i ($1 \leq i \leq m$). Each R_i $(1 \leq i \leq m)$ may coincide with the entire detector plane or may be some region in between. Each of the false signals $|\langle f_i,h_x\rangle|^2$ $(1 \leq i \leq m)$ should be small for all choices of x in $R_i - B_i$

$(1 \leq i \leq m)$. The first, most basic fundamental figure of merit used throughout this report is the signal–to–noise ratio of h, defined to be

$$SNR_1(h) = T_1(h)/N_1(h), \tag{1}$$

where

$$T_1(h) = \min_{1 \leq i \leq n} \max_{x \in B_i} |\langle f_i, h_x \rangle|^2 \tag{2}$$

and

$$N_1(h) = \max_{1 \leq i \leq m} \max_{x \in R_i - B_i} |\langle f_i, h_x \rangle|^2 \tag{3}$$

Here the subscript 1 is used on $SNR_1(h)$, $T_1(h)$, and $N_1(h)$ because in later sections other figures of merit will be considered in slightly different contexts. Call $T_1(h)$ the threshold of h and call $N_1(h)$ the noise of h. $T_1(h)$ is the largest threshold which will accept each of the images $f_1,\ldots,f_n$ as a true target. This definition makes sense, for it is merely the ratio of the smallest peak signal given by h off a true target to the largest false signal given by h. In order for h to be useful in a threshold detector, it is necessary that $SNR_1(h)$ be as large as possible (certainly greater than 1) while keeping $T_1(h)$ of reasonable size. SNR_1 is a slight generalization of the definition of signal–to–noise ratio introduced in Reference 8. In the case $n = m = 1$, SNR_1 is reminiscent of the notion of signal–to–clutter ratio used in radar (cf. Reference 9, p. 1–7). For any given discriminant function h, the number $SNR_1(h)$ depends on the choice of f_i's, B_i's, and R_i's, but this dependence is usually not explicitly noted when it is clear from context.

A discriminant function h gives rise to a (continuous)

phase-only filter provided each $\mathcal{F}(h)^*(k)$ is either a complex number of modulus 1 or, extending the notion of phase-only filter a bit, is 0. A discriminant function h gives rise to an n-state (discrete) phase-only filter provided each $\mathcal{F}(h)^*(k)$ is one of the nth roots of unity, $\exp(2\pi ik/n)$, $0 \leq k < n$. Of special interest in this report are those n-state phase-only filters for which n = 2 and n = 16. If n = 2, then each $\mathcal{F}(h)^*(k)$ is +1, -1, or 0, and $\mathcal{F}(h)^*$ is then termed a binary phase-only filter. n-state phase-only filters for which n = 2 or n = 16 are of much interest since they potentially can be implemented in addressable spatial light modulators (References 10 - 15). With such applications in mind, $\mathcal{F}(h)^*(k)$ usually is 0 for all k outside of some square box centered at origin (0,0) in the filter plane and is constant on subsquares corresponding to device pixels.

Next, what are the free parameters involved in the design of a good continuous phase-only filter? There are certain free parameters which involve the physical features of the optical correlator. Among these are the number of pixels on and/or the resolution of the input device, the number of pixels on and/or the resolution of the detector, the physical dimensions of the correlator, the characteristics of the lenses, and the wavelength of coherent radiation used. These are important engineering questions that probably will be resolved by device constraints and engineering tradeoffs determined by the particular application of the correlator. The proper choice of these free parameters is not addressed in this work. Another very important question in the optical correlator design is the proper choice of scaling in the Fourier or filter plane. In other words, what portion of the the filter plane should be intercepted or filtered by the spatial light modulator? Too small of a scaling will result in important spatial frequencies of potential target scenes being ignored, while too large of a scaling will result in

insufficient filtering being done, no matter how good the filter design. For an extreme case of the latter just consider the case in which one pixel of the spatial light modulator intercepts most of the important spatial frequencies of the input image. A happy medium, chosen over a wide range of input images, is necessary. A number of authors have written papers purporting to answer this scaling question (see Reference 16). These efforts seem to be misdirected and misleading. First of all, the design criteria they employ center around reducing circuit noise when in fact the chief problem in correlator design is discrimination of true target from false targets (clutter). Secondly, their design criteria are good only if there is a single image in the training set. Phase-only filters are notoriously sensitive to perturbations in the training set, so a training set of size one is far too small to be of any use. Thirdly, the answer is image dependent. This is clearly of little use for an addressable spatial light modulator which has a bank of thousands of different filters. The solution of the scaling problem in the filter plane probably cannot admit a theoretical solution but only an engineering solution which will be resolved at the point in time at which the overall correlator design is fixed. This report does not address this scaling problem but assumes it has been solved in a representative manner such as described in more detail in that portion of Section III concerned with the proper computer simulation of optical correlators.

Once the training set f_i, the regions R_i, and subregions B_i $(1 \leq i \leq m)$ of the filter plane have been fixed, the remaining free parameters are entries in the matrix representing the filter $\mathcal{F}(h)^*$, each of which is an independent parameter taking on the value zero or a complex number of modulus one. The choice of the f_i and the corresponding R_i and B_i $(1 \leq i \leq m)$ can only be determined by an engineering methodology involving much experimentation and computer simulation studies. In this study

the f_i's were supplied by the Air Force Armament Laboratory (AFATL). The R_i's usually were chosen to be either the detector face or the entire correlation plane, and the B_i's, after some experimentation, were usually chosen to be a relatively small box about origin (0,0) in the correlation plane. These choices are discussed in detail in the next section. Having fixed the f_i, R_i, and B_i, one should design h so that $\mathcal{F}(h)^*$ is a phase-only filter with $SNR_1(h)$ as large as possible. One therefore can construct an h with $\mathcal{F}(h)^*$ a phase-only filter and with $SNR_1(h)$ optimized only by varying these independent parameters.

Simulations reported in Reference 8 suggest that $\mathcal{F}(h)^*(k)$ should also be 0 for some k near and including the origin (0,0) in the filter plane in order for $SNR_1(h)$ to be large. In References 17 - 19 Flannery et al. give some decision rules for choosing which pixels in $\mathcal{F}(h)^*$ to set equal to 0 for binary phase-only filters, but such rules do not seem well suited for the large training sets to be considered later. It appears that an extrapolation of the work presented here can resolve in a rational, general manner the question of which pixels of $\mathcal{F}(h)^*$ to set equal to zero for training sets of arbitrary size, independent of whether the filter is continuous or discrete, and then give an optimal choice for the phases of the nonzero pixels. This study concentrated on what seemed to be the most challenging question, the proper choice of the pixels in $\mathcal{F}(h)^*$ which are of modulus one. The optimal choice of which pixels in $\mathcal{F}(h)^*$ to set equal zero was sidestepped. Instead, after a little experimentation, a quite arbitrary square set of low frequency pixels in $\mathcal{F}(h)^*$ were set equal to zero. Notice that there are an enormous number of free parameters in the optimization question of interest, for if $\mathcal{F}(h)^*$ is to be implemented in an $L \times L$ device, then there are L^2 free

parameters less the number of pixels set equal to 0. In this study L typically is 128, so L^2 is 16,384, while an 11 × 11 box, or 121, low frequency pixels were usually arbitrarily set equal to zero. Hence, 16,263 is a typical number of free parameters involved in phase-only filter design in this report.

Given that one has designed a good continuous phase-only filter, how should one go about discretizing it in a rational manner, as might be required for implementation in an actual device? An algorithm for doing this for binary phase-only filters was first given in Reference 8. That algorithm can easily be extended to the case of discrete n-state phase-only filters. It may be described geometrically as follows. Orient the circle in counter-clockwise fashion. Let z be a point on the unit circle greater than or equal to 1.0 but strictly less than $\exp(2\pi i/n)$. Moving counter-clockwise around the unit circle, let $A_k(z)$ be the arc of the unit circle greater than or equal to $\exp(2\pi ik/n)z$ but strictly less than $\exp(2\pi i(k+1)/n)z$, for each $0 \leq k < n$. Discretize $\mathcal{F}(h)^*$ into an n-state phase-only filter $\mathcal{F}(h_z)^*$ as follows. If $\mathcal{F}(h)^*[i,j]$ is the i,j-th entry of $\mathcal{F}(h)^*$ and $\mathcal{F}(h)^*[i,j] = 0$, then set $\mathcal{F}(h_z)^*[i,j] = 0$. Otherwise $\mathcal{F}(h)^*[i,j]$ is a complex number of modulus one and so lies on some unique arc $A_k(z)$. In this case let $\mathcal{F}(h_z)^*[i,j] = \exp(2\pi ki/n)$. Finally do a one-dimensional search in your favorite manner to find that $z = z_0$ so that $SNR_1(h_z)$ is as large as possible. This choice of z_0 then gives, at least in this sense, an optimal way to discretize $\mathcal{F}(h)^*$ into an n-state filter. A careful optimal choice of z_0 as prescribed can yield filters with more than twice the signal-to-noise ratio than a random choice of z when the training set is large. Moreover, especially for n = 16, the performance of the discretized phase-only filters faithfully mirrors that of their continuous phase-only ancestors.

SECTION III

TECHNICAL DETAILS OF THE IMAGERY AND SIMULATIONS

The Air Force Armament Laboratory supplied a great deal of the imagery used in this study. Most of the imagery was supplied on several VAX-compatible computer tapes in two series. Each series was produced by a visible light intensity digitizing camera from models on a target cloth background from essentially the same aspect and range over a swath of angles. The first series of target images, supplied in 1986, consisted of M113, M48, M1, and T62 images. For each target in the first series the swath of angles consisted of 21 views, the front flush view plus or minus 10 degrees at 1-degree increments, and an additional 21 views, the side flush view plus or minus 10 degrees with 1-degree increments. This gave a total of 168 images. The second series of images, supplied by the Air Force Armament Laboratory late in 1988, consisted of M1 and T62 images. For each target in the second series the swath of angles consisted of 61 views, the front flush view plus or minus 30 degrees with 1-degree increments, and an additional 61 views, the side flush view plus or minus 30 degrees with 1-degree increments. This gave a total of 244 images. Each image was 128 pixels by 128 pixels in size. The intensity of each pixel was scaled to be an integer between 0 and 255 and was stored in byte format. Additional images consisting of targets of strategic interest were furnished in 1988 and 1989. This report is not sequenced in chronological order of accomplishment or discovery but in the ex post facto most logical order of exposition. In each succeeding section the images used will be carefully identified as to type and series.

Each target in the two series of imagery needed to be separated from its background. To facilitate this the targets were uniformly brightly illuminated so that the background

intensity off the target cloth was digitized to be 255. The target then was simply extracted from the background by setting all pixels with intensity value 255 equal to 0. This process certainly worked well enough to provide sufficiently good clean target imagery, unencumbered by a background, for a conceptual study such as this. A coherent optical correlator utilizes the amplitudes of an image rather than its intensities (References 1 – 7). It therefore was necessary to convert the furnished intensity imagery to amplitude imagery for a correct computer simulation of a coherent optical correlator. To accomplish this for a clean target intensity image, each pixel was divided 255, the square root of each quotient was taken, the square root was multiplied by 255, and then this scaled square root was rounded to the nearest integer between 0 and 255. This gave a very good approximation to a clean target amplitude image unencumbered by a background. This final converted imagery was that used in simulation studies. Figure 2 shows a side view of an M48 from the first series of target images typical of the imagery used in this study. Note that the conversion from intensity imagery to amplitude imagery had a tendency to make each image somewhat flatter, and therefore only made it harder to discriminate one class of target from another.

The simulations to be described in the following sections were done in a manner to model reality as closely as possible. In computer simulations the entire plane first of all is replaced by a square centered at (0,0). The larger the square the better the approximation to the plane in the general neighborhood of (0,0). This square is made into a periodic pattern in the plane by identifying opposite edges. Such a periodic pattern in the plane can be viewed as one such pattern completely occupying the surface of a doughnut. Such a doughnut locally replicates the plane well at points not too remote from (0,0). This continuous doughnut is then replaced by a discrete doughnut periodic along each axis. If the discretization is fine enough, then the

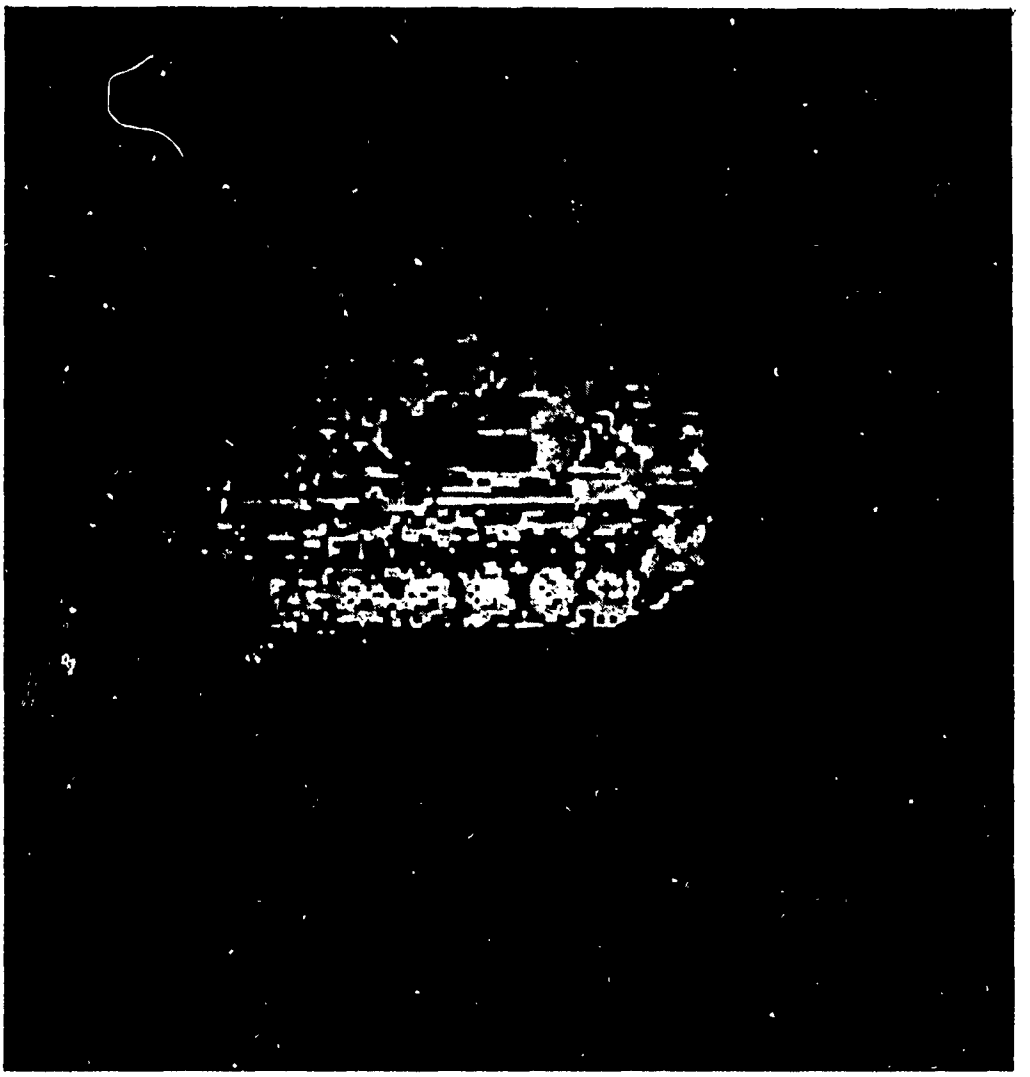


Figure 2. M48 Amplitude Image in Blank Background

discrete doughnut is itself a reasonable approximation to the plane at points not too remote from (0,0). The points on the doughnut usually are identified with the entries in a suitable two-dimensional matrix. Images on the plane are locally approximated by functions on this discrete doughnut, and usually represented in matrix form, two-dimensional integrals are replaced by doubly-indexed finite sums over two-dimensional matrix entries, and two-dimensional Fourier transforms are replaced by two-dimensional finite Fourier series. The latter are invariably evaluated using the two-dimensional fast Fourier transform algorithm.

One gross feature of a real correlator is that the input device is bounded in size. This must be taken into account in realistic simulations of the Fourier transforming properties of lenses. An examination of the derivation of the Fourier transforming properties of lenses shows that the paraxial assumption and the infinite lens approximation demand that the image being Fourier transformed certainly be smaller than the lens size. This elementary fact is quite important and, as will be seen in the next section, must be taken into account. As noted in previous paragraphs, all of the training set imagery was 128 × 128 in extent. The input device was modeled to be a 128 × 128 square, though any other shape (rectangular, circular, etc.) could have been modeled and would have led to similar problems and successes. To take into account that the input device is bounded and not the entire plane, the 128 × 128 imagery was inserted into the middle of a larger matrix which modeled the entire plane, which was always 256 × 256 in size. A square with an even number of rows and columns has no natural center. It was assumed that the center (0,0) of the input plane corresponded to the (129,129) entry of the 256 × 256 matrix. This forced the (65,65) entry of the 128 × 128 subimage representing the input device to correspond to (0,0) in the input plane. Similar precautions were taken in the Fourier plane since any real

spatial light modulator intercepts only a bounded portion of the Fourier plane. Namely, the spatial light modulator was assumed to be a 128 × 128 device which was placed in the low 128 × 128 frequency entries of the 256 × 256 models of the Fourier transform of the input plane. Thus only relatively low frequency terms in the Fourier plane were utilized. It is plausible that this choice made the recognition problem more challenging. Scaling the spatial light modulator to intercept a larger portion of the Fourier plane probably would have lead to better target recognition and discrimination results on the training sets, though probably at a cost of more instability in the recognition process for images close to but not included in the training sets. As mentioned in Section II, an arbitrary decision was made to set equal to zero the 11 × 11 box of low frequency pixels in the spatial light modulator. An optimal choice of which low frequency pixels to set equal to zero is deferred to another time.

SECTION IV

A BASIC DIFFICULTY IN TARGET DISCRIMINATION FOR PHASE-ONLY FILTERS OF STANDARD DESIGN

A simple but basic difficulty in the standard correlation process will now be pointed out. The author first noticed this phenomenon in 1984 – 1985, brought it to public notice in Reference 20, and gave another talk on it on November 3, 1988. First, the basic phase-only matched filter construction of Horner and Gianino will be recalled from Reference 21. Given a single image f, let $\mathcal{F}(h)^*$ be the phase-only filter defined by: $\mathcal{F}(h)^*(k) = 0$ if $\mathcal{F}(f)(k) = 0$ and $\mathcal{F}(h)^*(k) = \mathcal{F}(f)(k)^*/|\mathcal{F}(f)(k)|$ otherwise. For the purposes of this study the Horner-Gianino construction was altered slightly in conformity with the considerations of Section III. For example, if f is a 128 × 128 image, f was first embedded into the middle a 256 × 256 array of zeros to obtain a 256 × 256 image f', and a phase-only filter h was defined by: $\mathcal{F}(h)^*(k) = 0$ if $\mathcal{F}(f')(k) = 0$, or if k is inside the 11 × 11 box of low frequency pixels, or if k is outside the 128 × 128 box of low frequency pixels, and $\mathcal{F}(h)^*(k) = \mathcal{F}(f')(k)^*/|\mathcal{F}(f')(k)|$ otherwise. Here and in future sections, given an image f, the phase-only filter $\mathcal{F}(h)^*$ just constructed is termed the phase-only matched filter associated with f.

Some simple experiments were performed. Let t_1 be the M48 image in a blank background as illustrated in Figure 2 and let $\mathcal{F}(h_1)^*$ be the associated phase-only matched filter. In a simulation, t_1 was used as the input to an optical correlator with the phase-only filter $\mathcal{F}(h_1)^*$ in the Fourier plane and the resulting light intensity distribution in the correlation plane was computed. Figure 3 is a graph of the computed maximal light

Figure 3. t_1 vs. $\mathcal{F}(h_1)^*$, South-to-North Correlation Plane View

intensities on south–to–north lines over the detector in the correlation plane. The sharp spike in Figure 3 is at the center of the detector face, which in turn is centered in the correlation plane. The horizontal bar at the top of the spike is a reference level for the threshold of this filter. It will be used for comparison purposes in the next figure.

The filter $\mathcal{F}(h_1)^*$ did indeed recognize the M48. But it is rare to have a bright object in a zero intensity background. A new and more realistic image t_2 was designed which consisted of the image t_1 embedded in a uniform background whose amplitude was equal to the average amplitude of t_1. This new image t_2 is more realistic than the image t_1. This certainly is the case for infrared imagery, where objects tend to thermal equilibrium with their backgrounds. In a simulation, the new image t_2 was used as the input to an optical correlator with the phase–only filter $\mathcal{F}(h_1)^*$ in the Fourier plane and the resulting light intensity distribution in the correlation plane was computed. Figure 4 is a graph of the computed maximal light intensities on south–to–north lines over the detector in the correlation plane. There is no identification spike at all, for the threshold (represented by the horizontal bar at the top) to peak signal ratio is 42.82. This is in marked contrast to the correlation intensities for the classical (not phase–only) matched filter and, at first glance, is a bit of a surprise. This phenomenon holds independent of whether the input device is correctly modeled to be bounded in extent, as described in Section III, or whether the input device is incorrectly modeled to cover the entire input plane. It is true that some sort of edge enhancement technique probably could have been used to extract an edge enhanced version of the M48 from its background, but it is desirable to avoid this extraneous step if possible. There was no noise in the image t_2, so the observed phenomena cannot be attributed to noise. Large amounts of noise will be added to

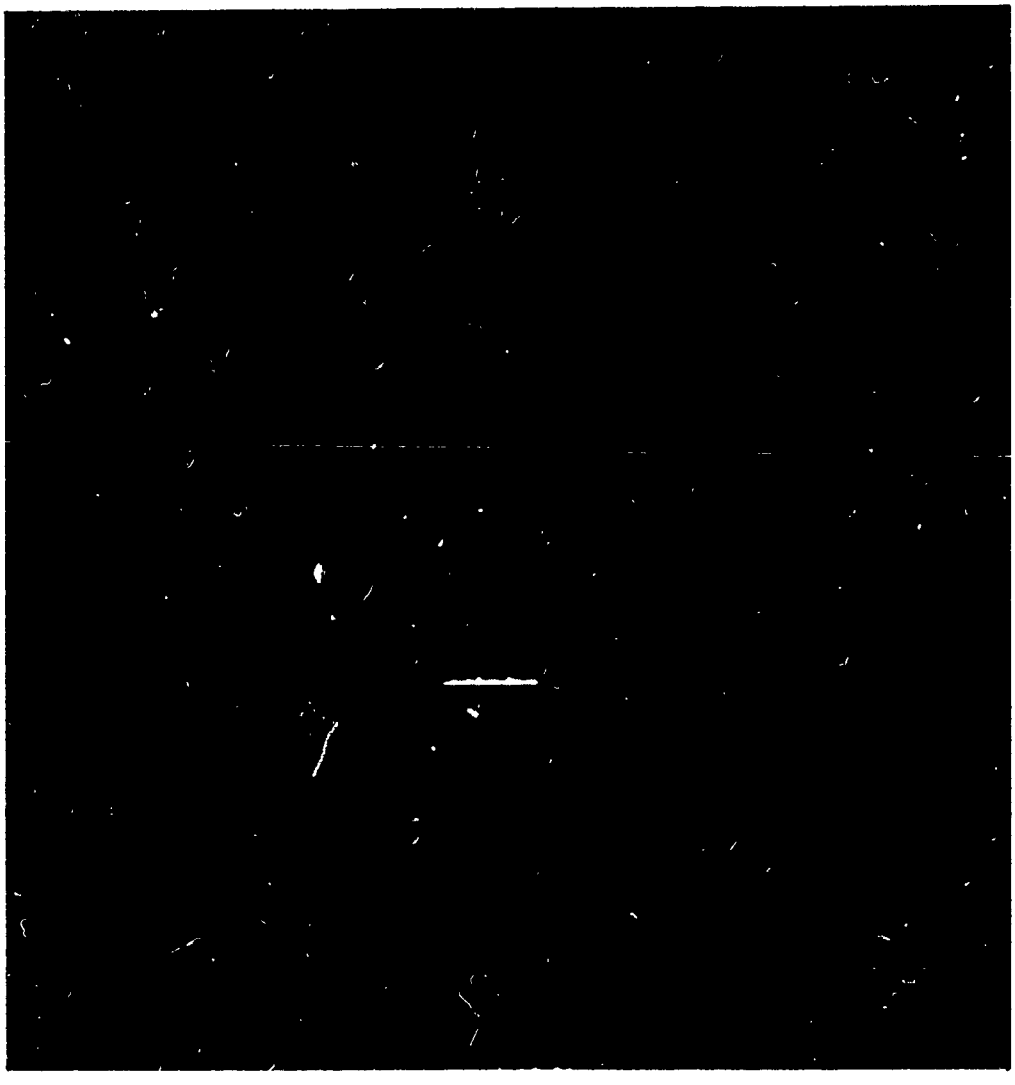


Figure 4. t_2 vs. $\mathcal{F}(h_1)^*$, South-to-North Correlation Plane View

images, similar to t_2, considered in later sections, and on these noisy images edge enhancement techniques probably would be useless for target identification and discrimination. The precise reason why this simulation gave its results defies simple explanation, because such an explanation is equivalent to an explanation for why the phase-only filtering process gives such tremendous discrimination. However, the following is a crude plausibility argument for the observed result of this last simulation. First of all, an elementary computation shows that if z, u, and v are complex numbers, if u is much greater than v and $z = u + v$, then the phase of z is close to the phase of u. This implies that if g is the sum of two functions g_1 and g_2, and if the g_1 has a much larger L^2-norm than does g_2, then the spatial frequency content of g is weighted most heavily at those frequencies where g_1 is most heavily weighted, and the phases of $\mathcal{F}(g)$ and $\mathcal{F}(g_1)$ will be close at those spatial frequencies. Hence, the associated phase-only matched filters of g and g_1 will give close to the same response to the input g. Now the dominant feature of t_2 is not the tank but simply the box. Hence, the associated phase-only matched filter of t_2 will be much more like that of the box and not that of the tank t_1, at least at the most energetic spatial frequencies of t_2. Hence, one would expect the responses of the associated phase-only matched filters for t_1 and t_2 to be quite different, as indeed they are.

Another experiment was conducted. Let $\mathcal{F}(h_2)^*$ be the phase-only matched filter associated to the image t_2. In a simulation, the image t_2 was input to an optical correlator with the phase-only filter $\mathcal{F}(h_2)^*$ in the Fourier plane and the resulting light intensity distribution in the correlation plane was computed. Figure 5 is a graph of the computed maximal light intensities on south-to-north lines over the detector in the correlation plane. The horizontal bar again is an artificial construct representing a threshold and will be used for

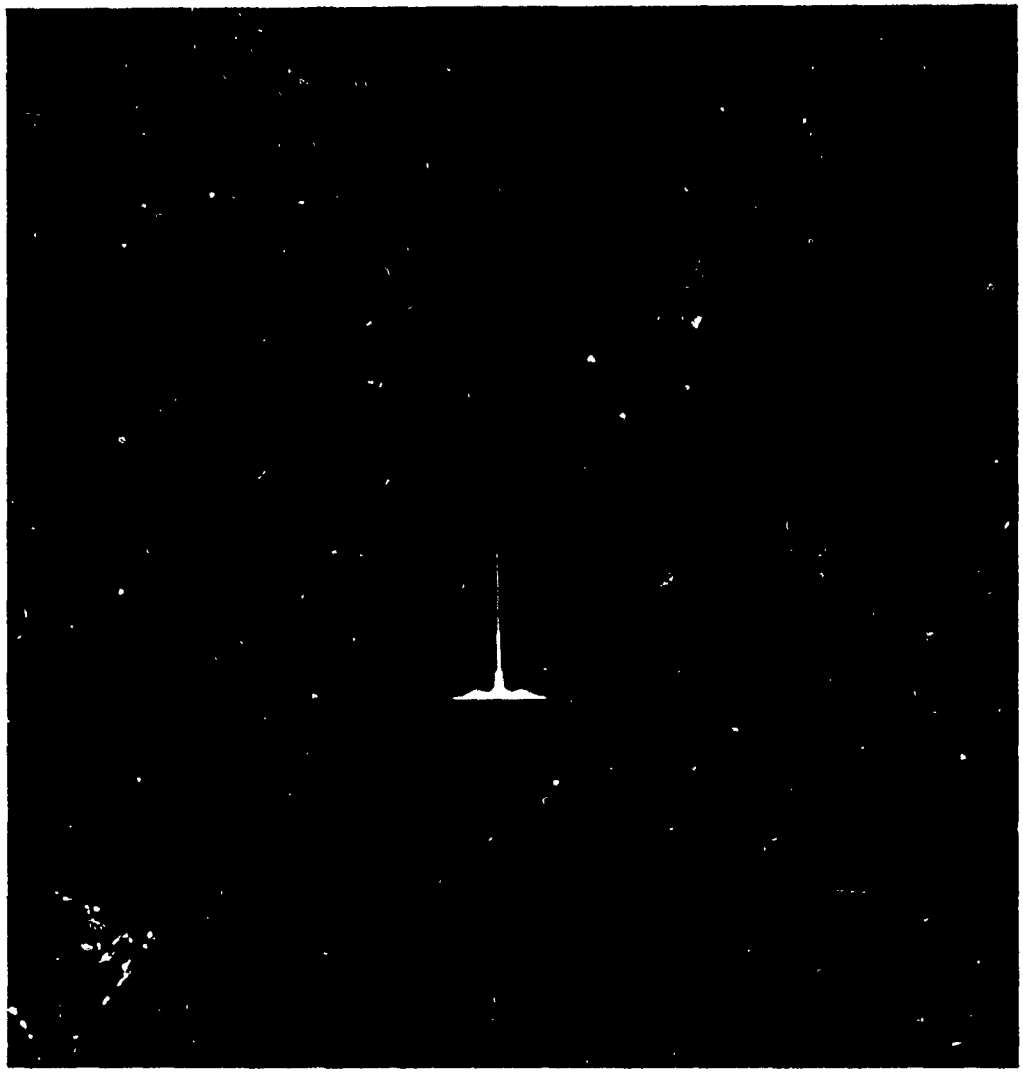


Figure 5. t_2 vs. $\mathcal{F}(h_2)^*$, South–to–North Correlation Plane View

comparison purposes in the next figure. There is an initial difficulty with the filter $\mathcal{F}(h_2)^*$, for it cannot be translation invariant, in the sense that if the M48 moves about in the constant background in the box representing the input device, then the correlation spike given by the filter $\mathcal{F}(h_2)^*$ will not follow its movement. The filter $\mathcal{F}(h_2)^*$ would track the movement of the M48 only if the box representing the constant background moved along with it, which is manifestly not the case. While this difficulty must be kept in mind, it is not the chief one at the moment. Consider the simplest of all possible false targets which might have some chance of confusing $\mathcal{F}(h_2)^*$. Namely, let t_3 be the image whose amplitude is constant and is the same as that of the background of t_2. In a simulation, the image t_3 was used as the input to an optical correlator with the phase-only filter $\mathcal{F}(h_2)^*$ in the Fourier plane and the resulting light intensity distribution in the correlation plane was computed. Figure 6 is a graph of the computed maximal light intensities on south-to-north lines over the detector in the correlation plane. The horizontal bar again represents the threshold of Figure 5. The discrimination in this most simple case is unsatisfactory, for the ratio of the threshold to the peak signal is only 2.32.

A plausibility argument for this very poor discrimination is easily given and is equivalent to the crude plausibility argument given previously. What is the overwhelmingly prominent feature of t_2? It is the box filled with the average background, namely t_3. It is simple to check that if two images are very similar or close, then their Fourier transforms are close on average, and so the phases of their Fourier transforms are close on average. Hence, the phase-only matched filter associated with t_2 should be close to the phase-only matched filter associated with t_3. Therefore the correlation plane light intensity distribution produced by an optical correlator with the image t_2 in the input

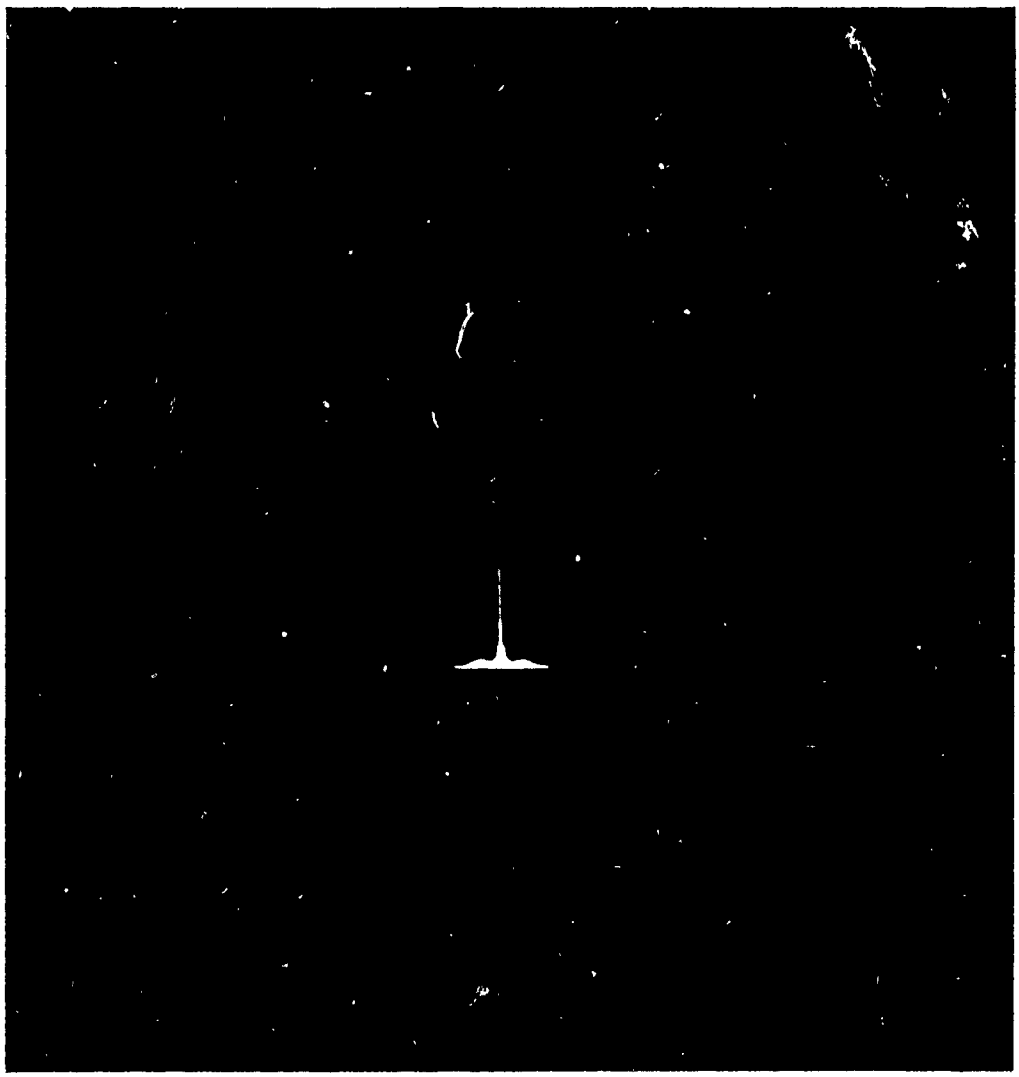


Figure 6. t_3 vs. $\mathcal{F}(h_2)^*$, South–to–North Correlation Plane View

plane and with $\mathcal{F}(h_2)^*$ in the filter plane should in large measure be the same as the correlation plane light intensity distribution produced by the optical correlator with the image t_3 in the input plane and with $\mathcal{F}(h_2)^*$ in the filter plane.

This phenomenon apparently has not been reported in the literature. There may be several explanations for this. First, in actual equipment demonstrations of optical correlation using phase-only filters many researchers use a two-state zero-one input device, so that each pixel in the input device is either turned on or not. Objects of interest embedded in nontrivial backgrounds cannot be used as inputs to an optical correlator with such devices. Second, in published examples of equipment demonstrations of optical correlation using phase-only filters which are not limited to two-state inputs, either a nontrivial background was not present, or, if a background indeed was present, the demonstration was limited to recognition and not discrimination. The possible reason that these phenomena were not discovered in computer simulations can be attributed to several things. First, perhaps only two-state zero-one input imagery was used. Secondly, and perhaps more likely, the correlations were not done correctly in the sense that the input device was not modeled to be bounded. If the computer modeling is not done correctly and the constant background is identified with the entire torus and therefore is not bounded in extent, then the background effect is not noticed. To see this, take a constant function c on a two-dimensional discrete torus. The Fourier transform of c is a constant multiple of the Dirac delta function at (0,0) in the Fourier transform torus. Then any phase-only filter whose (0,0) entry is 0 will give no correlation with the constant background.

SECTION V

THE OBJECTIVES OF THE PROGRAM

The discussions in Section II and Section IV strongly suggest the importance of solving the following two problems.

Problem #1: Develop and implement in computer codes efficient and effective algorithms for phase-only filter design which fully utilize all of the independent phase parameters.

Problem #2: Use such algorithms to design phase-only filters whose discretizations reliably extract many aspects of a target of interest from nontrivial backgrounds in a noise immune and translation invariant manner while simultaneously discriminating against false targets.

The positive solution of these two problems was the major theme of this work. These two problems in large measure have been solved.

Concerning Problem #1, at first glance it might not seem feasible to directly find an optimal or near optimal set of phases which maximize objective functions such at $SNR_1(h)$ or $T_1(h)$. After all $\mathcal{F}(h)^*$ is an $L \times L$ array where L typically is 128, 256, or larger, and the number of free parameters involved in the optimization of $SNR_1(h)$ will be L^2 minus the number of low frequency entries set equal to 0. Such an optimization problem involves far too many variables to be solved by exhaustive search or random search methods. Furthermore, $SNR_1(h)$ is nondifferentiable and so does not possess gradients. Nonetheless, in the course of this study an efficient and effective class of iterative algorithms have been devised which, when carefully implemented into computer codes, give very good and perhaps even globally optimal solutions. Results based on the first generation of such algorithms and computer codes were

reported in References 22 and 23. It suffices to make a crude guess for the initial point of the iterative process. Such a guess can be generated by choosing a subset of $f_1, \ldots, f_n$ and applying the techniques of Horner and Gianino (References 21 and 24), as modified in a minor way along the lines discussed in Section IV. These iterative algorithms and computer codes appear to be relatively insensitive to the choice of initial point, for near optimal filters were obtained in several numerical experiments starting from a randomly generated phase-only filter. An outline of the ideas involved in these iterative algorithms is given in the Appendix A. If $n = m = 1$ and a phase-only filter for the standard correlation architecture is to be constructed, then these iterative algorithms essentially reproduce the phase-only matched filter of Horner and Gianino (Reference 21). However, if $m > 1$, these iterative algorithms differ radically from the prescription of Horner and Gianino (Reference 24) and are necessary for optimal or near optimal filter design (References 22 and 23). These iterative algorithms and computer codes are of considerable general utility, for, as will be seen in later sections, they are also necessary for optimal phase-only filter design for the new filtering concepts to be introduced in later sections. A major advantage of these iterative algorithms and computer codes is that no sets of linear equations need be solved and no matrix inverses need to be computed once an initial guess is generated, so that it is feasible to optimize $SNR_1(h)$ using a very large number of images. Since phase-only filters are notoriously sensitive to small perturbations in the training set, this ability to design a filter which can be optimized over a very large number of images is useful. Simple variants of these iterative algorithms can be used to drive $T_1(h)$ to a maximum or $N_1(h)$ to a minimum. In some instances, in particular for two of the new filtering concepts to be introduced later, simply maximizing $T_1(h)$ results in a significant increase in $SNR_1(h)$, for $N_1(h)$ usually undergoes minor oscillations about its

initial value. This approach gives some control over the smallest peak signal from true targets. In the two other filtering concepts discussed later, quite surprisingly, simply optimizing $SNR_1(h)$ yields a great increase in $SNR_1(h)$ while giving up very little of the size of $T_1(h)$ in the worst case and usually yielding a significant increase in $T_1(h)$. This is important, for while it is in general impossible to simultaneously maximize $SNR_1(h)$ and $T_1(h)$, light budget considerations imply that one does not want to maximize $SNR_1(h)$ at the expense of driving $T_1(h)$ to a low value. The first generation of algorithms and computer codes used in this study in the construction of phase-only filters had extremely long running times on a VAX 11-785, which frankly was very badly configured (many users, a low priority, and very limited memory necessitating a great deal of page faulting and disk I-O). However, during the course of this project, the efficiency of the algorithms and codes has progressed steadily and considerably, so that currently all of the optimization problems described in some detail later can now probably be solved on the same VAX 11-785 in several hours of CPU time. The author and others, independently, estimate that typical running times on a state-of-the art dedicated workstation would be a half hour or less. Since the official ending date of this project some versions of the design codes have been drastically upgraded and rewritten in the language occam 2. These occam 2 codes have been run on some quite inexpensive INMOS T800-25 transputer modules installed on boards in an IBM-AT located in the author's office. The transputer and the language occam 2 were designed together. Programming in occam 2 permits one to carefully optimize code to run on transputers. Initial test runs indicate that two to five minutes suffice for the design of good phase-only filters in this new computing environment. INMOS also claims that transputer modules will be available early in 1991 which have five to 10 times the floating point processing power of the T800-25

(Reference 25). Thus one does not need a supercomputer for the approach to phase-only filter design advocated in this report. Objections to the methods of phase-only filter design advocated in this report which claim that such methods are too computationally intensive to be useful thus seem irrelevant and groundless. Such is true in view of the availability of inexpensive current hardware, the speed at which the design codes appear to run on these current inexpensive devices, and the potential availability of much more powerful inexpensive hardware in the near future. After all, it is the performance of the filters which is most important, not that they were designed using only a VAX 11-785, or a PC-XT, or a TI-30 pocket calculator, or an abacus, or a slide rule, or just pencil and paper.

Concerning Problem #2, it was pointed out in Section IV that the average background amplitude level in which a target might be embedded must be considered when designing phase-only filters. It seems inescapable that different filters are needed when the same target is embedded in different amplitude backgrounds. The number of filters needed to extract a particular target from the spectrum of all possible reasonable backgrounds is not addressed in this work. Rather this study concentrated on the case for which there was the least contrast, for this seemed to be the most challenging case. Very simple estimates prove that the correlation plane response of $\mathcal{F}(h)^*$, and therefore $SNR_1(h)$, varies continuously if the background into which an input image is embedded varies smoothly. Preliminary work has shown that the filters to be described in this study are quite robust over a fairly wide range of backgrounds centered around the average amplitude of the target training sets. Some degradation of performance is noted as the size of the training sets is increased. A thorough study of questions in this area need to conducted in a systematic manner.

The basic philosophy or intuitive idea adopted in this study in attempts to solve Problem #2 is the following: most of the important information in an image is contained in its variations from its average amplitude. This intuitive idea is discussed at length in References 26 – 28. The idea behind this solution philosophy can be expressed mathematically as follows. Let f be an image which vanishes outside of an aperture A, let χ_A be the indicator function of A ($\chi_A(x) = 1$ if $x \in A$, $\chi_A(x) = 0$ if $x \notin A$), and let E(f) be the expectation of f or the average of f over A. It is elementary that $f = E(f)\chi_A + (f - E(f)\chi_A)$ and $E(f - E(f)\chi_A) = 0$. The summand $f - E(f)\chi_A$ contains most of the important information content of f for purposes of recognition and discrimination. It also is relatively insensitive to changes in the background. Rather than use f, $f - E(f)\chi_A$ should be used. Somehow, either mathematically, algorithmically, or architecturally, when confronted with the image f seen through the aperture A, a good target recognition and discrimination system should subtract the average amplitude $E(f)\chi_A$ from f.

The first rather naive implementation of this idea was to make phase–only filters by the method of Horner and Gianino (Reference 24), as slightly modified along the lines discussed in Section IV, using the training set images $f_i - E(f_i)\chi_A$ $(1 \leq i \leq m)$ and not the training set images f_i $(1 \leq i \leq m)$. This change led to some improvement in filter performance, but it failed to give a good solution even in the very special case for which n = m = 1. For example, let h_3 be the discriminant function such that $\mathcal{F}(h_3)^*$ is the phase–only matched filter associated with the image $t_2 - E(t_2)\chi_A$. Here t_2 is the M48 image of Figure 2 in a background in a square aperture, as discussed in Section IV. In a simulation, the image t_2 was used as the input to an optical correlator with the phase–only filter $\mathcal{F}(h_3)^*$ in the Fourier plane and the resulting light intensity distribution in the correlation plane was computed. Figure 7 is a graph of the

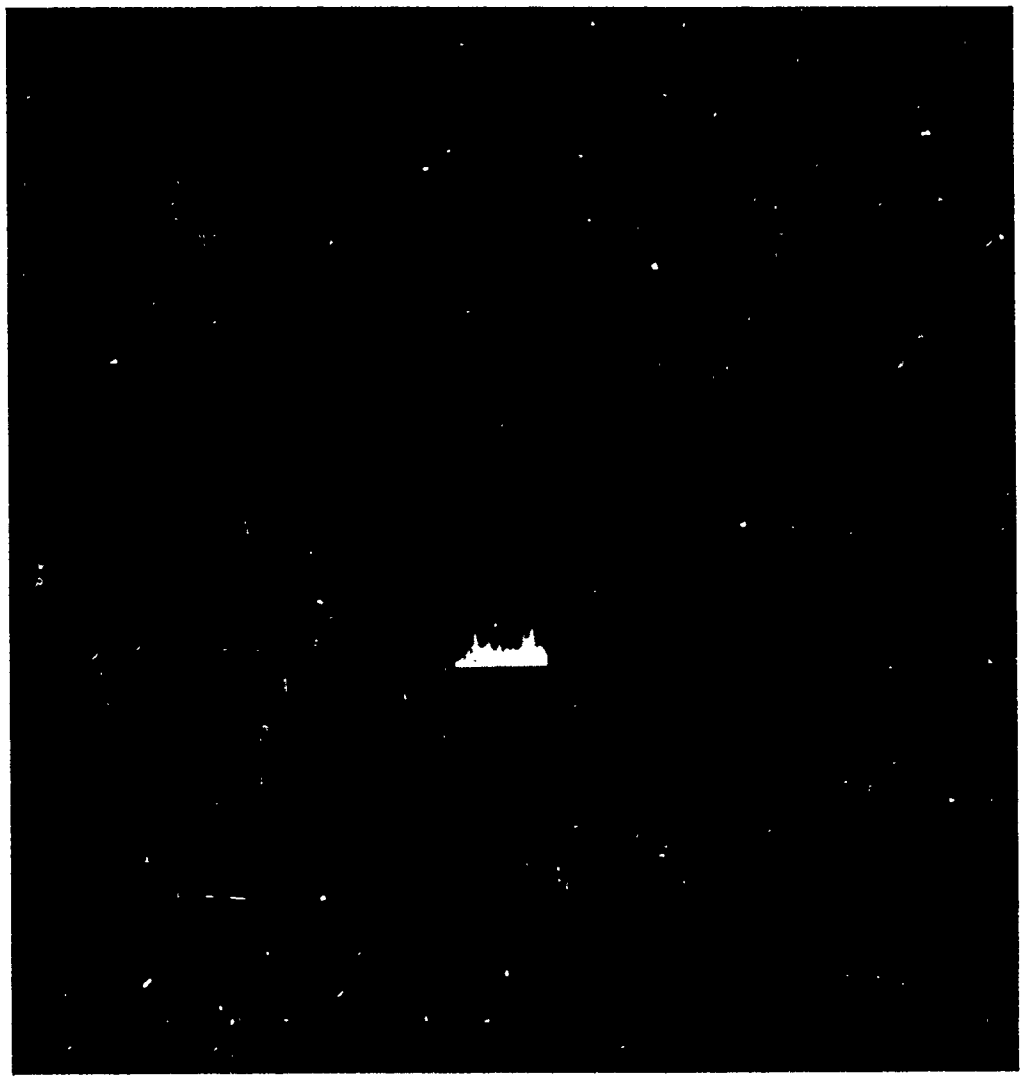


Figure 7. t_2 vs. $\mathcal{F}(h_3)^*$, South–to–North Correlation Plane View

computed maximal light intensities on south–to–north lines over the detector in the correlation plane. The horizontal bar at the top, as usual, is an artificial construct representing a threshold and will be used for comparison purposes in the next figure. The filter $\mathcal{F}(h_3)^*$ definitely recognized the target t_2. Next, as in Section IV, let t_3 be the simplest false target, namely let t_3 be the image whose amplitude is constant and is the same as that of the background for t_2. Another simulation was performed, the image t_3 was used as the input to an optical correlator with $\mathcal{F}(h_3)^*$ in the Fourier plane, the corresponding light intensity distribution in the correlation plane was computed, and the maximal light intensities along south–to–north lines over the detector was displayed in Figure 8. A comparison of Figure 8 with Figure 6 shows that the use of the filter $\mathcal{F}(h_3)^*$ instead of the filter $\mathcal{F}(h_2)^*$ led to some improvement in discrimination of t_2 from t_3, for in Figure 8 the threshold to peak signal level has improved to 5.6 for t_3 versus $\mathcal{F}(h_3)^*$, whereas in Figure 6 the threshold to peak signal level for t_3 versus $\mathcal{F}(h_2)^*$ was 2.32. There is some improvement in filter performance, but the improvement is not dramatic enough to give hope that this simple modification of the method of Horner and Gianino can be extended to very large training sets. Four different techniques for drastic improvements in filter performance are given in the next four sections.

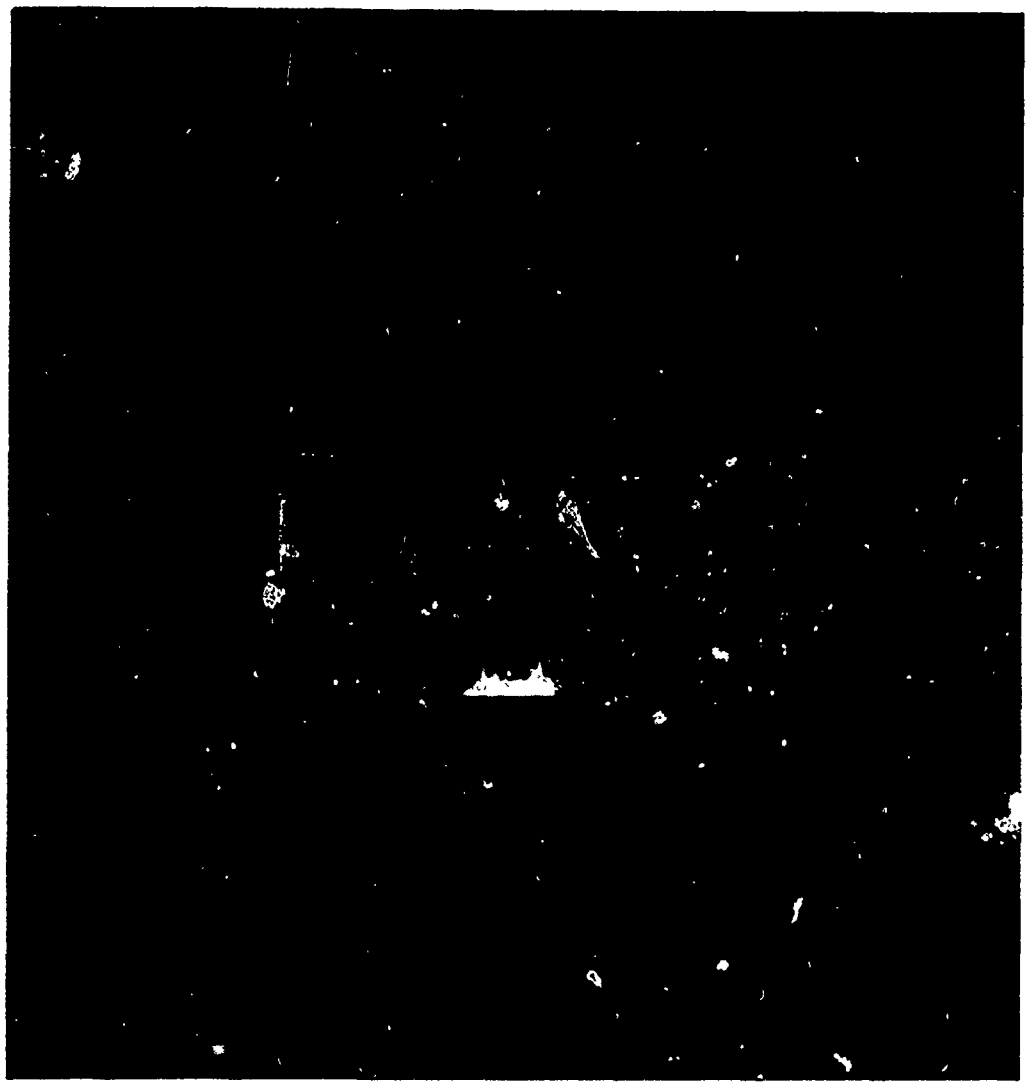


Figure 8. t_3 vs. $\mathcal{F}(h_3)^*$, South-to-North Correlation Plane View

SECTION VI

THE FIRST SOLUTION: PROPER FILTER DESIGN

The first solution to Problem #2 raised in Section V is totally straightforward. Namely, using the notation introduced in Section II, given a set of true targets $g_1, \ldots, g_n$ in a blank background (such as a subset of the second series of 61 T62 amplitude images discussed in Section III), compute their average nonzero pixel value C, place each target image g_i in a square aperture with C as the constant background amplitude to obtain a new true target image f_i, let $m = n + 1$, and set f_{n+1} to be the false target image which is just the square aperture with C as the constant background amplitude. Then starting with the phase-only filter discussed at the end of Section V for the training set $f_1, \ldots, f_n, f_{n+1}$, use the algorithms discussed in Section V and Appendix A to construct a discriminant function h so that $\mathcal{F}(h)^*$ is a phase-only filter with $SNR_1(h)$ optimized as best as possible. Do this optimization carefully as discussed in Section III.

Notice that by suppressing the peak signal off f_{n+1} not only has there been an improvement in the ability of the phase-only filter $\mathcal{F}(h)^*$ to discriminate against f_{n+1}, but also the performance of $\mathcal{F}(h)^*$ has been made much more robust with respect to translations of any g_i $(1 \leq i \leq n)$, as long as the image g_i stays wholly within the field of view in the constant amplitude background C. To see this, note that sliding g_i around inside the aperture with a constant amplitude background of C is precisely the same as sliding $f_i - C\chi_A$ around inside the aperture with a constant amplitude background of C. More precisely, let T be any rigid translation so that $T(g_i)$ lies wholly inside the aperture. If the background in which g_i resides has a constant

amplitude of C, then the new image in the aperture coincides with $T(f_i - C\chi_A) + C\chi_A$.

But

$$|\langle[T(f_i - C\chi_A) + C\chi_A], T(h)\rangle|^2$$
$$= |\langle T(f_i - C\chi_A), T(h)\rangle|^2 + 2\mathrm{Re}(\langle T(f_i - C\chi_A), T(h)\rangle \cdot \langle C\chi_A, T(h)\rangle) + C^2|\langle \chi_A, T(h)\rangle|^2$$
$$= |\langle (f_i - C\chi_A), h\rangle|^2 + 2\mathrm{Re}(\langle (f_i - C\chi_A), h\rangle \cdot \langle C\chi_A, T(h)\rangle) + C^2|\langle \chi_A, T(h)\rangle|^2$$

since T preserves inner products, and

$$|\langle f_i, h\rangle|^2$$
$$= |\langle[(f_i - C\chi_A) + C\chi_A], h\rangle|^2$$
$$= |\langle (f_i - C\chi_A), h\rangle|^2 + 2\mathrm{Re}(\langle (f_i - C\chi_A), h\rangle \cdot \langle C\chi_A, h\rangle) + C^2|\langle \chi_A, h\rangle|^2.$$

Hence,

$$||\langle[T(f_i - C\chi_A) + C\chi_A], T(h)\rangle|^2 - |\langle f_i, h\rangle|^2|$$
$$= |2\mathrm{Re}(\langle (f_i - C\chi_A), h\rangle \cdot \langle C\chi_A, T(h)\rangle) + C^2|\langle \chi_A, T(h)\rangle|^2 - 2\mathrm{Re}(\langle (f_i - C\chi_A), h\rangle \cdot \langle C\chi_A, h\rangle) - C^2|\langle \chi_A, h\rangle|^2|$$
$$\leq 2|\langle (f_i - C\chi_A), h\rangle| \cdot |\langle C\chi_A, T(h)\rangle| + 2|\langle (f_i - C\chi_A), h\rangle| \cdot |\langle C\chi_A, h\rangle| + |\langle C\chi_A, T(h)\rangle|^2 + |\langle C\chi_A, h\rangle|^2$$
$$\leq 2|\langle f_i, h\rangle| \cdot |\langle C\chi_A, T(h)\rangle| + 2|\langle C\chi_A, h\rangle| \cdot |\langle C\chi_A, T(h)\rangle| + 2|\langle f_i, h\rangle| \cdot |\langle C\chi_A, h\rangle| + 2|\langle C\chi_A, h\rangle|^2 + |\langle C\chi_A, T(h)\rangle|^2 + |\langle C\chi_A, h\rangle|^2.$$

Next, divide this string of inequalities by $|\langle f_i, h\rangle|^2$ to obtain

$$|\{|\langle[T(f_i - C\chi_A) + C\chi_A], T(h)\rangle|^2 / |\langle f_i, h\rangle|^2\} - 1|$$
$$\leq 2(|\langle C\chi_A, T(h)\rangle| / |\langle f_i, h\rangle|) + 2\{(|\langle C\chi_A, h\rangle| / |\langle f_i, h\rangle|) \cdot (|\langle C\chi_A, T(h)\rangle| / |\langle f_i, h\rangle|) + 2(|\langle C\chi_A, h\rangle| / |\langle f_i, h\rangle|) + 2(|\langle C\chi_A, h\rangle| / |\langle f_i, h\rangle|)^2 + (|\langle C\chi_A, T(h)\rangle| / |\langle f_i, h\rangle|)^2 + (|\langle C\chi_A, h\rangle| / |\langle f_i, h\rangle|)^2.$$

Let $R = 1/\mathrm{SQRT}(\mathrm{SNR}_1(h))$. Then

$$|\{|<[T(f_i - C\chi_A) + C\chi_A], T(h)>|^2/|<f_i,h>|^2\} - 1|$$
$$\leq 2R + 2R^2 + 2R + 2R^2 + R^2 + R^2$$
$$= 2R(2 + 3R).$$

Thus, the ratio

$$|<[T(f_i - C\chi_A) + C\chi_A], T(h)>|^2/|<f_i,h>|^2$$

can be driven close to one by driving $SNR_1(h)$ to a high level. In other words, the larger $SNR_1(h)$ is made, the more translation invariant is the optical correlation process. This statement is true under the assumption that C is fixed. A perturbation of the above argument gives simple but crude estimates of the translation invariance of the optical correlation process when both C and T vary independently and, of course, shows continuity in both parameters. The bounds which one derives to prove these statements probably are not particularly sharp. Empirical evidence gathered from the rather large imagery used in this study suggests that the optical correlation process using phase-only filters of this study's design probably is much more translation invariant than these simple but crude bounds suggest. This supposition needs careful checking.

Tests of the scheme proposed in the first paragraph of this section have been carried out on a variety of subsets of the second series 61 T62 images discussed in Section III. Some of these results are presented in Table 1 – Table 5. During the course of these calculations the filters were purposely optimized so that the recognition spike for each true target occurred at (0,0) in the correlation plane, each phase-only filter had its low 11 × 11 frequency terms set equal to 0, each of R_i $(1 \leq i \leq n + 1)$ was chosen to be the correlation plane or, perhaps more interestingly, was chosen to coincide with the detector face, $B_i = R_i$ $(1 \leq i \leq n)$, and $B_{n+1} = \phi$.

Table 1 shows the performance rating for various continuous phase-only filters before and after optimization via the iterative design techniques of this study. In this table the

TABLE 1. SNR_1 FOR PHASE-ONLY FILTERS MADE FROM SUBSETS OF THE 61 T62 IMAGES, OPTIMIZED OVER THE ENTIRE CORRELATION PLANE

Number of Tanks	SNR_1 Initial Filter	SNR_1 Optimal Filter
1	71268.56/19982.37 = 3.566572	110779.0/1779.360 = 62.25776
37	11269.12/22391.76 = 0.5032709	24373.51/1648.890 = 14.78177
41	8821.926/21648.14 = 0.4075142	22668.08/1682.985 = 13.46898
45	5010.160/20405.42 = 0.2455309	21424.40/1738.687 = 12.32217
49	4944.087/19661.24 = 0.2514637	19747.55/1634.902 = 12.07874
53	4087.492/20220.60 = 0.2021449	19380.76/1559.984 = 12.42369
57	4090.192/21021.73 = 0.1945697	19125.26/1699.251 = 11.25511
61	5640.813/21439.39 = 0.2631054	18251.88/1675.026 = 10.89647

TABLE 2. SNR_1 FOR PHASE-ONLY FILTERS MADE FROM SUBSETS OF THE 61 T62 IMAGES, OPTIMIZED OVER THE DETECTOR FACE

Number of Tanks	SNR_1 Initial Filter	SNR_1 Optimal Filter
1	71268.56/19982.37 = 3.566572	87470.72/228.2725 = 383.1855
5	46572.79/19619.90 = 2.371203	62066.14/263.6993 = 262.2152
9	32132.33/16839.45 = 1.908158	44300.57/362.0043 = 122.3758
13	25778.24/16752.72 = 1.538750	32797.94/270.5114 = 121.2442
17	17080.28/17463.89 = 0.9780338	27203.71/270.0250 = 100.7452
21	13256.70/16731.74 = 0.7923086	21924.45/253.7263 = 86.40987
25	10639.40/18885.06 = 0.5633767	19875.38/256.5071 = 77.48471
29	11132.16/20728.79 = 0.5370384	18490.65/225.5611 = 81.97627
33	9865.263/22054.50 = 0.4473129	17805.41/284.2386 = 62.64246

TABLE 2. SNR_1 FOR PHASE-ONLY FILTERS MADE FROM SUBSETS OF THE 61 T62 IMAGES, OPTIMIZED OVER THE DETECTOR FACE (CONCLUDED)

Number of Tanks	SNR_1 Initial Filter	SNR_1 Optimal Filter
37	11269.12/22391.76 = 0.5032709	16462.35/217.5071 = 75.68647
41	8821.926/21648.14 = 0.4075142	15170.02/240.6803 = 63.02972
45	5010.160/20405.42 = 0.2455309	14085.55/250.8399 = 56.15355
49	4944.087/19661.24 = 0.2514637	13078.30/225.5571 = 57.98221
53	4087.492/20220.60 = 0.2021449	12600.86/222.1102 = 56.73247
57	4090.192/21021.73 = 0.1945697	12373.76/235.2423 = 52.59580
61	5640.813/21439.39 = 0.2631054	11674.34/242.0804 = 48.73109

TABLE 3. PEAK SIGNALS OFF THE 61 T62 IMAGES GIVEN BY THE PHASE-ONLY FILTER LISTED IN TABLE 2 WITH THESE IMAGES AS THE TRAINING SET

T62 Image	Signal Intensity at (0,0) in the Correlation Plane
1	16916.76
2	15170.15
3	13452.24
4	11787.68
5	11698.94
6	12537.65
7	14182.69
8	12544.26
9	12237.69
10	12658.64
11	12227.11
12	11733.69
13	13207.46
14	13614.12
15	14852.85
16	13519.06
17	14907.09
18	14528.01
19	15235.75
20	14993.74
21	12383.46
22	13973.41
23	15186.59
24	13194.58
25	12686.87

TABLE 3. PEAK SIGNALS OFF THE 61 T62 IMAGES GIVEN BY THE PHASE-ONLY FILTER LISTED IN TABLE 2 WITH THESE IMAGES AS THE TRAINING SET (CONTINUED)

T62 Image	Signal Intensity at (0,0) in the Correlation Plane
26	14960.36
27	14099.18
28	11925.42
29	12239.57
30	11925.88
31	12869.59
32	13793.01
33	12867.13
34	11917.15
35	13178.68
36	12966.50
37	11747.31
38	12610.26
39	14202.35
40	12589.57
41	12779.90
42	15669.91
43	13808.50
44	13357.23
45	13579.48
46	17786.11
47	14196.47
48	11763.91
49	13652.38
50	12722.38

TABLE 3. PEAK SIGNALS OFF THE 61 T62 IMAGES GIVEN BY THE PHASE-ONLY FILTER LISTED IN TABLE 2 WITH THESE IMAGES AS THE TRAINING SET (CONCLUDED)

T62 Image	Signal Intensity at (0,0) in the Correlation Plane
51	12343.76
52	12088.06
53	11674.34
54	12356.88
55	11794.07
56	12338.33
57	15049.25
58	12134.26
59	13696.92
60	13710.62
61	15459.23

TABLE 4. SNR_1 FOR THE OPTIMIZED BINARY PHASE-ONLY FILTERS MADE FROM OPTIMIZED PHASE-ONLY FILTERS SELECTED FROM TABLE 1 AND TABLE 2

Number of Tanks	SNR_1 Optimized Binary Phase-only Filter
1	58647.23/3826.177 = 15.24819
21	12300.36/3850.205 = 3.194728
41	13865.34/3621.077 = 3.829066
53	10080.72/3449.044 = 2.922759
57	10792.47/3397.299 = 3.176779
61	11501.86/3231.800 = 3.558964

TABLE 5. SNR_1 FOR THE OPTIMIZED 16-STATE PHASE-ONLY FILTERS MADE FROM OPTIMIZED PHASE-ONLY FILTERS SELECTED FROM TABLE 1

Number of Tanks	SNR_1 Optimized 16-State Phase-only Filter
1	85631.95/368.2507 = 232.5371
21	21734.90/415.3951 = 52.3244
41	14766.46/389.8145 = 37.88073
53	13397.48/372.1299 = 36.0215
57	12698.56/415.6658 = 30.54992
61	11864.56/389.5699 = 30.45554

filters were designed by choosing each R_i $(1 \leq i \leq n + 1)$, in particular R_{n+1}, equal to the entire correlation plane in the optimization process. Each training set consisted of an odd number, 2k + 1 say, of the T62 images, the flush side view plus and minus k degrees at one degree increments. The initial filter was always taken to be that made from all 2k + 1 images in the training set by the method of Horner and Gianino (Reference 24), as trivially modified along the lines suggested in Section IV. $SNR_1(h)$ for each filter h is given as the quotient of two numbers which are the values of $T_1(h)$ and $N_1(h)$ in the units being used. The actual values of the numerator and denominator in each case are not particularly important since they are subject to drastic scaling in a real device, but the ratios of the various numerators and the various denominators indicate how $T_1(h)$ and $N_1(h)$ vary relative to one another as the number of training set images is changed. Two things are noteworthy. First, the optimization process of this study led to a 30- to 40-fold increase in $SNR_1(h)$. Secondly, this increase in $SNR_1(h)$ was not accompanied by a decrease in the threshold $T_1(h)$, which invariably increased after optimization by a factor of three or four when the training set became large.

Table 2 is similar to Table 1 but contains data for a larger collection of training sets. Furthermore, in this table the filters were designed by choosing each R_i $(1 \leq i \leq n + 1)$, especially R_{n+1}, to be the detector face in the correlation plane. Again, the training set imagery used consisted of subsets of the second series of 61 T62 side images discussed in detail in Section III, each subset consisted of an odd number, 2k + 1 say, of the T62 images, the flush side view plus and minus k degrees at 1 degree increments, and the initial filter was always taken to be that made from all 2k + 1 images in the training set by the method of Horner and Gianino (Reference 24), as trivially modified along the lines suggested in Section IV. In every instance $SNR_1(h)$ increased approximately two orders of magnitude

after optimization while the threshold $T_1(h)$ never decreased and most often increased significantly. There are a few anomalies in Table 2. Notice that SNR_1 for a filter made from 29 images is greater than that made from 25. This is a reflection of the fact that the algorithms of this study drive $SNR_1(h)$ to a local and not necessarily a global optimum. Notice however, that this was done at a cost decreasing the threshold $T_1(h)$. A similar anomaly exists between the optimized filters and thresholds for 33 and 37 images, and for 45 and 49 images. Notice that as the size of the training set decreased, the threshold monotonely decreased. There is evidence to believe that the most current versions of the phase-only filter design algorithms are better at finding global optima than the versions used to calculate Table 2. These anomalies probably would not occur if the optimizations were carried out using them. This assertion should be checked.

Table 3 illustrates the relative stability of the recognition signal coming off each of the 61 T62 images given by that optimized phase-only filter listed in Table 2 which had the 61 T62 images as part of its training set.

Table 4 contains the results of the one-dimensional optimization process described in Section II used to binarize selected optimized continuous phase-only filters from Table 2. These results are mildly disappointing. Table 5 contains the results of the one-dimensional optimization process described in Section II to discretize selected optimized continuous phase-only filters from Table 2 into 16 states. These results are much more pleasant. It is an interesting question as to why there is the large difference in performance between 2-state and 16-state filters. No satisfactory answer is known except to point to the details of the correlation process itself as used for discrimination. Binary filters seem to play a useful role in target recognition and even in discrimination when the image which is the false target is very unlike the training set images (see References 22 and 23). However, a two-state filter seems to

be inadequate for picking out relative low energy differences between two images whose most energetic spatial frequencies are very much alike, even when the very lowest frequency terms are suppressed, as they are here. The discussion of Section II implies that the target discrimination process using phase-only filters boils down to the following problem. Given a large number M of finite sequences of complex numbers of some fixed length k $\{ a(i,j)\ (1 \leq i \leq k) \mid 1 \leq j \leq M \}$, choose a fixed finite sequence of k complex numbers of modulus one, $z(1),\ldots,$ $z(k)$, so that

$$\max_{1 \leq j \leq M} \left| \sum_{1 \leq i \leq k} z(i)a(i,j) \right|^2$$

is as small as possible. It is clear that the more constraints which are put on the choice of the $z(i)$'s, the larger this minimum must be in general. Table 4 suggests that for the problem at hand constraining the $z(i)$'s to be +1 or −1 is too restrictive to give a good solution to this minimax question, but that constraining the $z(i)$'s to be chosen from the 16th roots of unity is not excessively restrictive. It is possible that a bad choice has been made for some of the gross parameters involved in filter design for binary phase-only filters to play a useful role. For example, rescaling the filter to intercept a larger or smaller portion of the Fourier plane, or suppressing a larger portion of the low frequency terms in the filter plane, or implementing a variant and/or splicing of the ideas of Flannery et al. and/or the ideas of this study might give phase-only filters which binarize in an acceptable manner for these large training sets. These options should be investigated. In any case, the data presented in Tables 4 and 5 certainly suggest the vast superiority of 16-state phase-only filters over binary phase-only filters, for they yield a 10-fold or more increase in SNR_i and in the size of useful training sets while requiring only four times the storage requirements. This conclusion tentatively

holds for the sort of optical correlation system discussed in this section. However, it will be shown that binary phase-only filters can still play a useful role for the systems described in Sections VII and VIII and especially for the filtering scheme described in Section IX.

It is of interest to test how well these 16-state discretized filters work. Let t_4 be the flush side view of a T62 in an aperture in a background whose amplitude is the same as the average amplitude of nonzero pixels of the original 61 T62 images in the second series of images. The image t_4 is displayed in Figure 9. Let $\mathcal{F}(h_4)^*$ be the optimized 16-state discretized filter made from the 61 T62 images in an aperture in a background whose amplitude is constant and equal to the average nonzero amplitude on the 61 T62 images. $SNR_1(h_4)$ is listed in Table 5. An experiment was conducted. In a simulation, the image t_4 was used as the input to an optical correlator with the phase-only filter $\mathcal{F}(h_4)^*$ in the Fourier plane and the resulting light intensity distribution in the correlation plane was computed. Figure 10 is a graph of the computed maximal light intensities on south-to-north lines over the detector in the correlation plane. The horizontal bar, as usual is an artifact representing $T_1(h_4)$ and will be used for comparison purposes in some of the next few figures. Let t_5 be the simple false target which consists of the aperture filled with a constant background whose amplitude is the same as the average amplitude of the nonzero pixels of the original 61 T62 training set images. In a simulation, the image t_5 was used as the input to an optical correlator with the filter $\mathcal{F}(h_4)^*$ in the Fourier plane and the resulting light intensity distribution in the correlation plane was computed. Figure 11 is a graph of the maximal light intensities along south-to-north lines over the detector in the correlation plane. The threshold to peak signal ratio is 30.46. Thus $\mathcal{F}(h_4)^*$ performs rather well

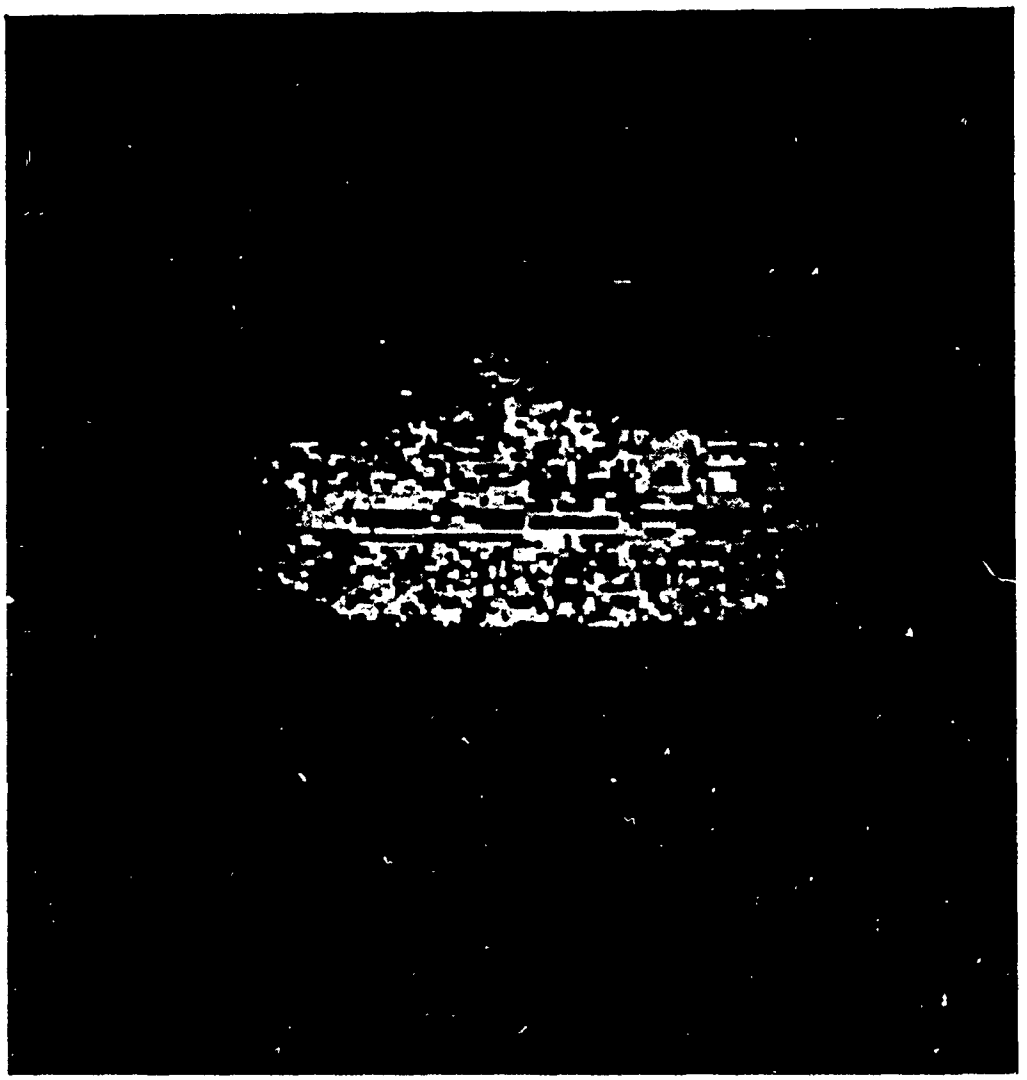


Figure 9. T62 Amplitude Image in a Nontrivial Constant Amplitude Background as Seen Through an Aperture

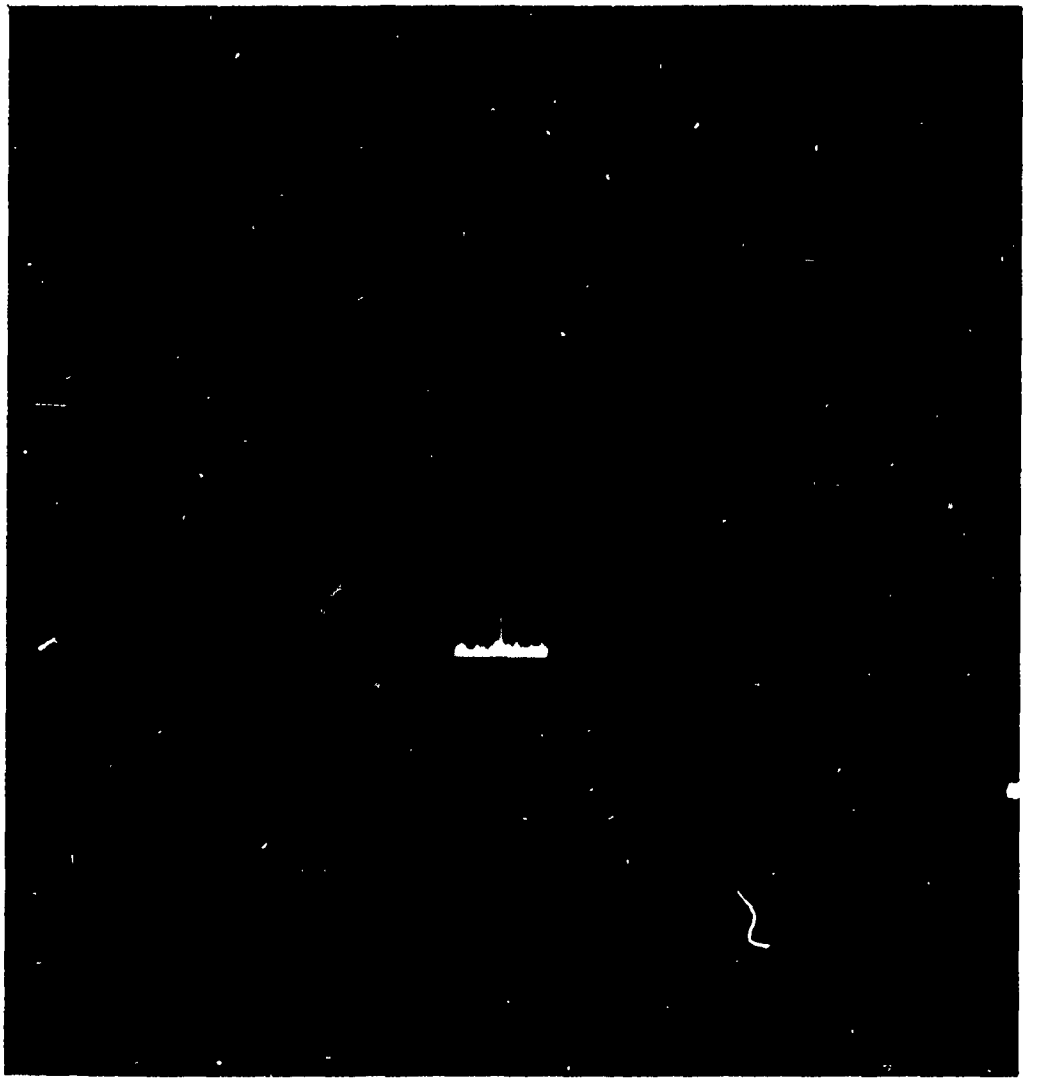


Figure 10. t_4 vs. $\mathcal{F}(h_4)^*$, South-to-North Correlation Plane View

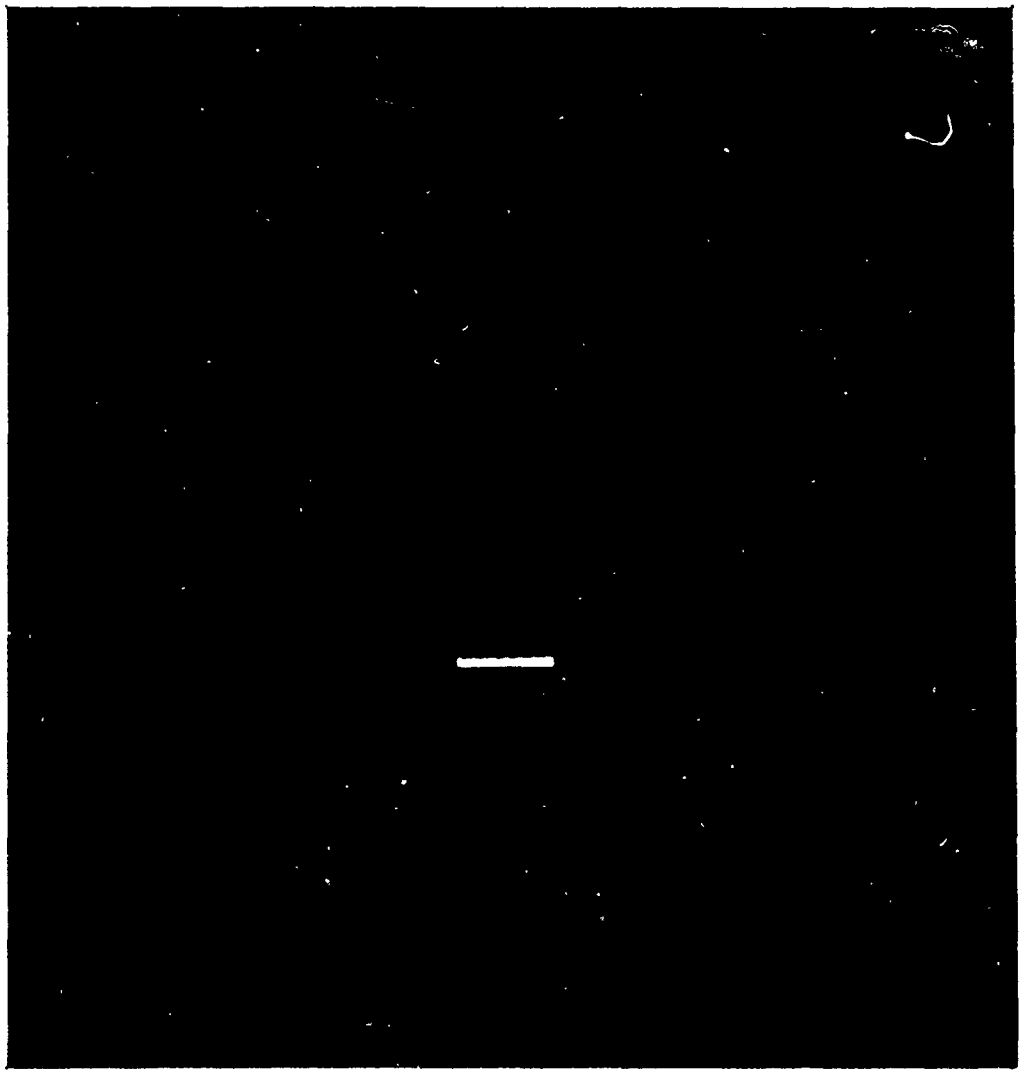


Figure 11. t_5 vs. $\mathcal{F}(h_4)^*$, South–to–North Correlation Plane View

in this discrimination test for which it has been designed. It is of interest to test the performance of $\mathcal{F}(h_4)^*$ against false targets which are not in the training set but which are vaguely similar to the T62 training set images. Let t_6 be the flush side view of an M1 in an aperture in a background whose amplitude is the same as the average amplitude of the nonzero pixels of the original 61 T62 images. The image t_6, taken from the second series of M1 imagery, is displayed in Figure 12. In a simulation, the image t_6 was used as the input to an optical correlator with the filter $\mathcal{F}(h_4)^*$ in the Fourier plane and the resulting light intensity distribution in the correlation plane was computed. Figure 13 is a graph of the resulting maximal light intensities along south–to–north lines over the detector in the correlation plane. The threshold to peak signal ratio is 5.94. This is typical of the performance of $\mathcal{F}(h_4)^*$ versus any of the M1 side images. Recall that this discrimination performance is achieved with no knowledge of the M1 images being used in the design of $\mathcal{F}(h_4)^*$. It is of interest to test the performance of $\mathcal{F}(h_4)^*$ against the training set images when they are buried in noise. Let t_7 be the image displayed in Figure 14. This image t_7 is the same as t_4 except that Gaussian noise with $\mu = 0.0$ and $\sigma = 25.0$ has been independently added to each pixel. In a simulation, the image t_7 was used as the input to an optical correlator with the filter $\mathcal{F}(h_4)^*$ in the Fourier plane and the resulting light intensity distribution in the correlation plane was computed. Figure 15 is a graph of the resulting maximal light intensities along south–to–north lines over the detector in the correlation plane. Target recognition is certainly achieved in spite of the noise. This is typical for the performance of $\mathcal{F}(h_4)^*$ against any of the T62 training set images with the same amount of noise added.

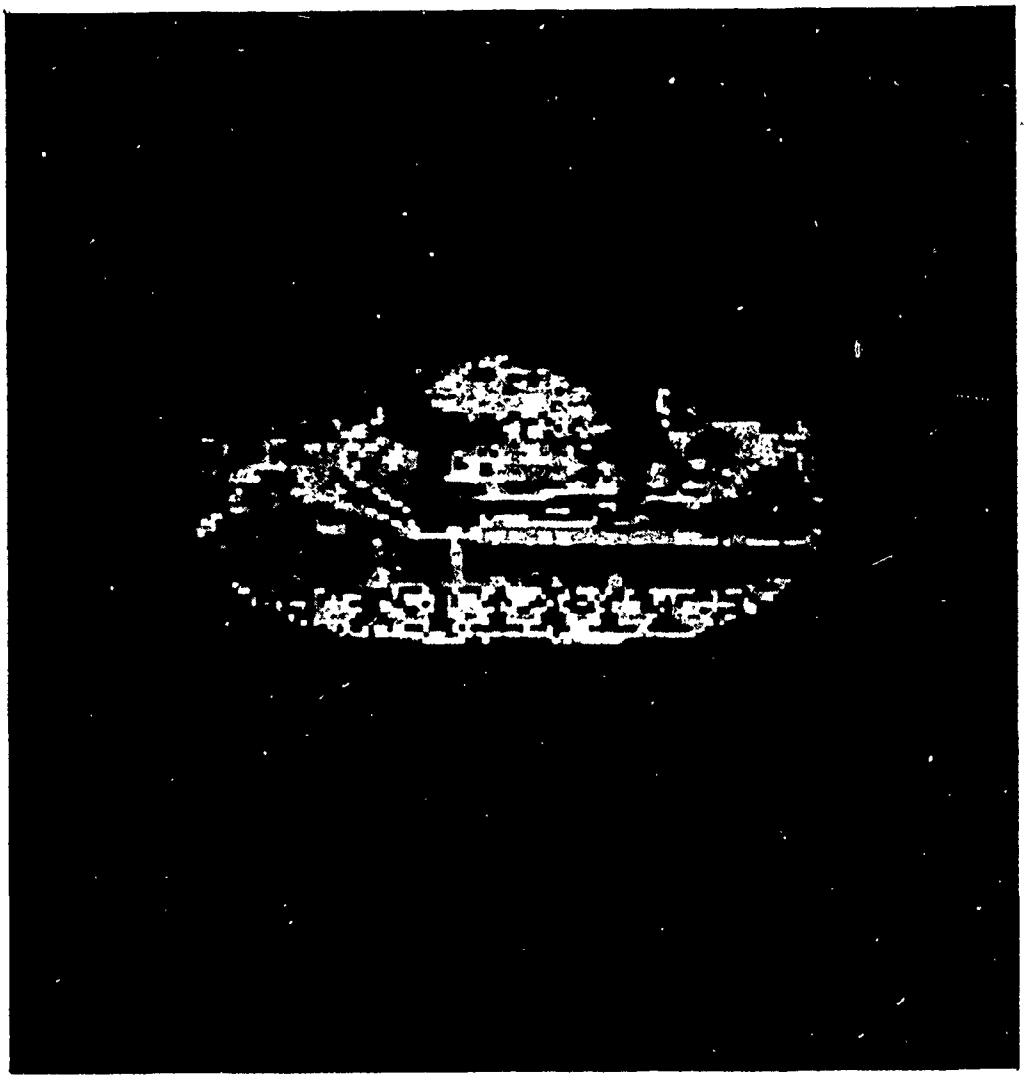


Figure 12. M1 Amplitude Image in a Nontrivial Constant Amplitude Background as Seen Through an Aperture

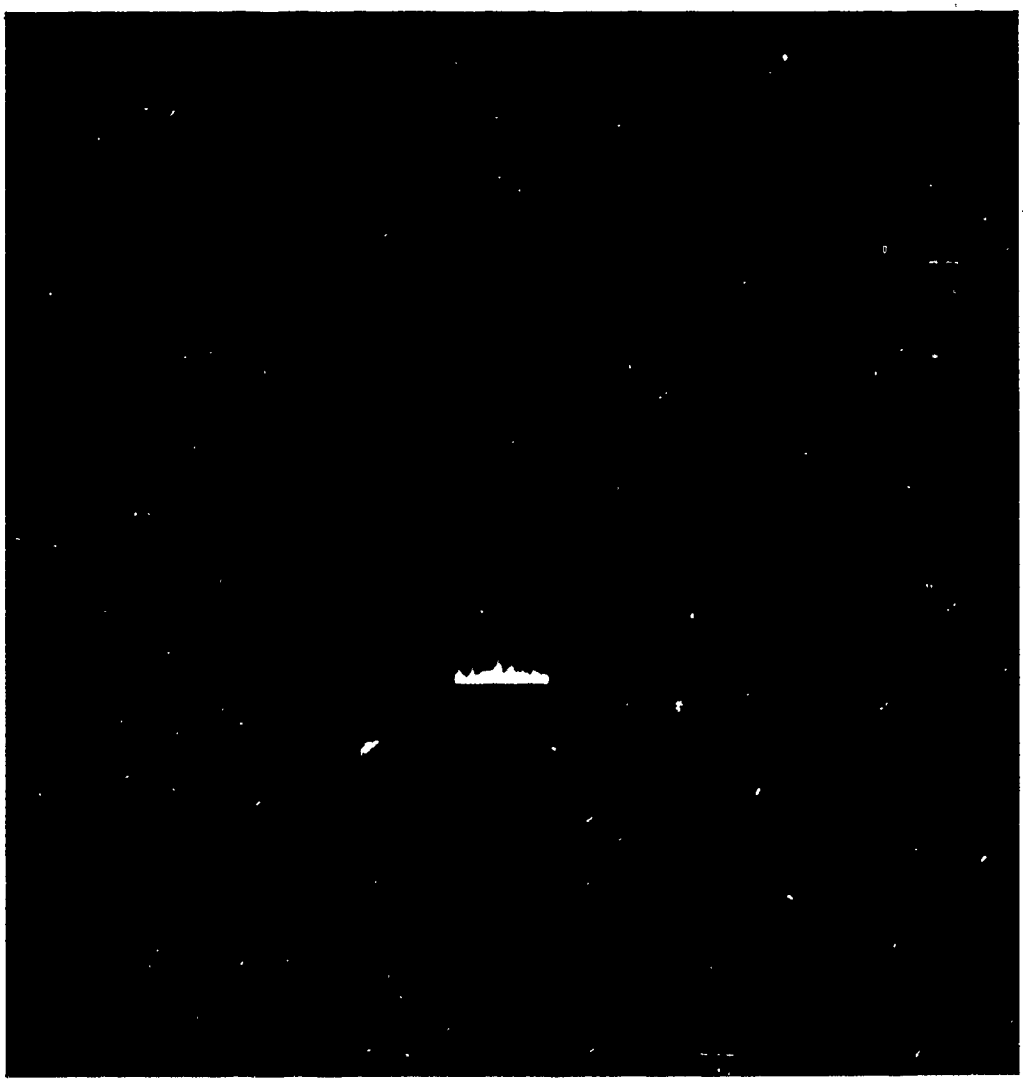


Figure 13. t_6 vs. $\mathcal{F}(h_4)^*$, South–to–North Correlation Plane View

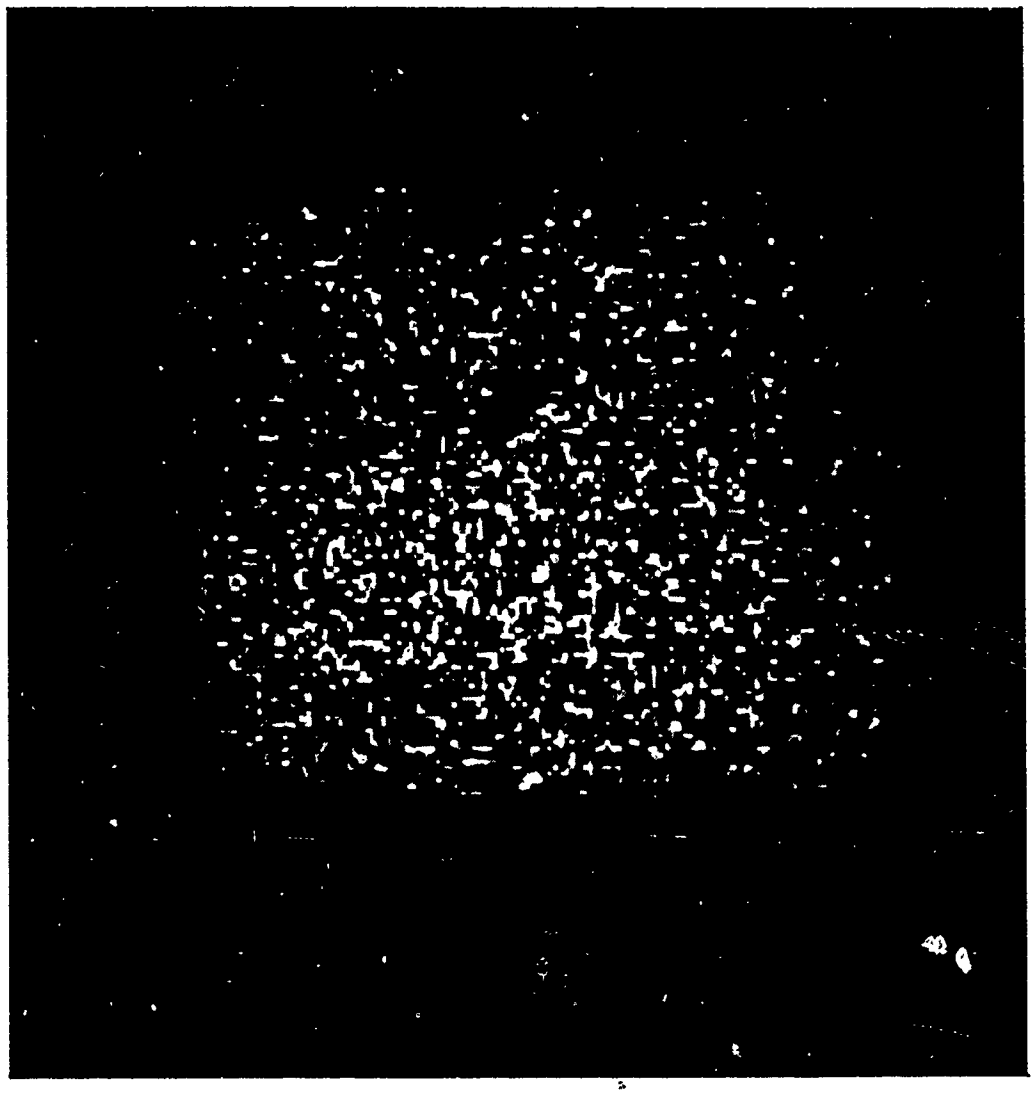


Figure 14. The T62 Image of Figure 9 Buried in μ = 0.0, σ = 25.0 Additive Gaussian Noise

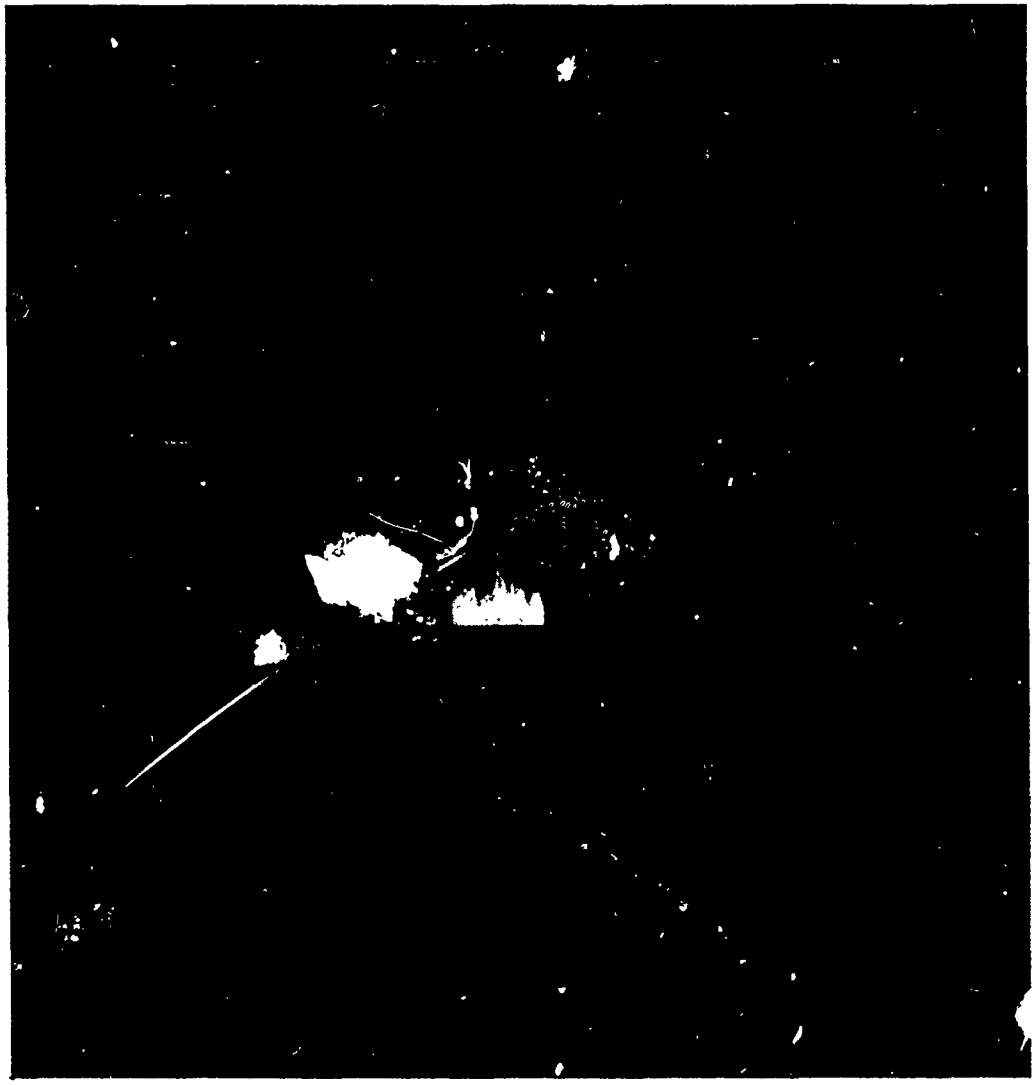


Figure 15. t_7 vs. $\mathcal{F}(h_4)^*$, South-to-North Correlation Plane View

SECTION VII

THE SECOND SOUTION: CODING INTENSITY IMAGES AS PHASE-ONLY IMAGES FOR OPICAL CORRELATOR INPUT PLUS PROPER FILTER DESIGN

The results of this section are of necessity quite preliminary since the ideas involved only occurred to the AFATL Program Manager and the author in a recent telephone conversation. Suppose a detector registers an intensity image $[a_{pq}^2]$ ($a_{pq} \geq 0$, $1 \leq p, q \leq N$) on its active pixels. In the standard optical correlator architecture and algorithm the amplitude image $[a_{pq}]$ is then used as the input to the correlator. There is no mathematical reason why $[a_{pq}]$ should be used as the input to the correlator. It might be advantageous to input some other image $[b_{pq}]$ which is a function of $[a_{pq}^2]$ to the correlator, so long as there is a one-to-one correspondence $[a_{pq}^2] \rightarrow [b_{pq}]$ between the detected image and the inputs. After all, there is a definite but subtle difference between the actual measured pixel values a_{pq}^2 and the information content they convey. If the measured intensity image $[a_{pq}^2]$ contains an object of interest, then the correspondence $[a_{pq}^2] \rightarrow [b_{pq}]$ should commute with translations of the object of interest inside the aperture represented by the matrix of detected intensities. This can be the case in general only if there is a function $\varphi : [0,+\infty> \rightarrow \mathbb{C}^1$ so that $b_{pq} = \varphi(a_{pq}^2)$. Here $\mathbb{C}^1$ is the set of complex numbers. This rule $[a_{pq}^2] \rightarrow [\varphi(a_{pq}^2)]$ can be one-to-one in general only if φ is one-to-one, as least on the set of possible

measured intensities a_{pq}^2. There is a wide body of folklore which states that amplitudes do not contain much information and are quite troublesome, but that phases do contain a great deal of information, for signals in general and for Fourier transforms of imagery in particular (Reference 29). This philosophy is suggestive but does not strictly apply here, for translation does not commute with the Fourier transform operator. Proceeding in a quite simple-minded fashion, however, this philosophy suggests taking φ to be a complex exponential. In the particular case at hand the intensities are circumscribed to be the integers between 0 and 255. This suggests letting $b_{pq} = \varphi(a_{pq}^2) = \exp(a_{pq}^2 \pi i/255.0)$. Here a_{pq}^2 is divided by 255.0 and then multiply by π so that $a_{pq}^2 \pi/255.0$ ranges between 0 and π. We do this so that $a_{pq}^2 \to \varphi(a_{pq}^2)$ is one-to-one and so that $\varphi(0)$ and $\varphi(255)$ are as far apart as possible. It is easy to think of more general functions φ which have similar properties. Note that in the case at hand $\varphi(a_{ij}^2)$ always will be a complex number of modulus one lying in the first or second quadrants of the complex plane.

There are devices described in the literature which might be used to input an arbitrary array of phases lying in the first or second quadrant of the complex plane to an optical correlator. For example, a reading of Reference 30 shows that the Hughes liquid crystal light valve can be used in such a mode. One may glean from Reference 31 that one can even use an inexpensive liquid crystal television for this purpose. It is somewhat ironic that the authors of Reference 31 were trying to operate their liquid crystal television as a full continuous phase spatial light modulator and appear to be somewhat vexed that they could only achieve a continuous phase response for phases lying

in the first and second quadrants.

It is of interest to test if there is an advantage to using phase-only inputs instead of the usual amplitude inputs to an optical correlator. A few of the numerical experiments described in Section VI were repeated to test if there might indeed be some advantage in their use. In these calculations intensity images coded as phase-only images, as described in the previous paragraph, were used as the inputs to a simulated optical correlator rather than the usual amplitude input images used in Section VI. The results are described in Table 6 – Table 11. These results, while very preliminary, are quite encouraging for this new approach to optical correlation since they appear to consistently produce results which are over four times that produced by the standard techniques. Let $\mathcal{F}(h_5)^*$ be the 16-state phase-only filter listed in Table 11 made from the optimized continuous phase-only filter whose training set included the 61 T62 side images, coded as phase-only inputs. In a simulation, the T62 image t_4 was coded into a phase-only image and used as the input to an optical correlator with the filter $\mathcal{F}(h_5)^*$ in the Fourier plane and the resulting intensity distribution in the correlation plane was computed. Figure 16 is a graph of the resulting maximal light intensities along south-to-north lines over the detector in the correlation plane. There is excellent recognition. The horizontal bar as usual is an artifact representing $T_1(h_5)$ and will be used for comparison purposes in the next few figures. Next, the simple false target image t_5 described in Section VI was coded into a phase-only image by the technique of this section. The previous numerical experiment was repeated with t_5 in place of t_4. The results are displayed in Figure 17, which indicate excellent discrimination. These last two results are to be expected from the statistics listed in Table 11. It is of interest to test the performance of $\mathcal{F}(h_5)^*$ against false targets for which it has no knowledge but which are

TABLE 6. SNR_1 FOR PHASE-ONLY FILTERS MADE FOR PHASE-ONLY INPUTS FROM SUBSETS OF THE 61 T62 IMAGES, OPTIMIZED OVER THE ENTIRE CORRELATION PLANE

Number of Tanks	SNR_1 Initial Filter	SNR_1 Optimal Filter
1	10.69354/0.4961187 = 21.55441	11.25012/0.043723952 = 257.2989
61	0.9933575/0.5285769 = 1.879306	1.652262/0.042049207 = 39.29354

TABLE 7. SNR_1 FOR PHASE-ONLY FILTERS MADE FOR PHASE-ONLY INPUTS FROM SUBSETS OF THE 61 T62 IMAGES, OPTIMIZED OVER THE DETECTOR FACE

Number of Tanks	SNR_1 Initial Filter	SNR_1 Optimal Filter
1	10.69354/0.4961187 = 21.55441	9.914534/0.006309473 = 1571.372
61	0.9933575/0.5285769 = 1.879306	1.421531/0.006232505 = 228.0834

TABLE 8. SNR_1 FOR THE OPTIMIZED BINARY PHASE–ONLY FILTERS MADE FROM THE OPTIMIZED PHASE–ONLY FILTERS LISTED IN TABLE 6

Number of Tanks	SNR_1 Optimized Binary Phase–only Filter
1	4.652245/0.090415299 = 51.45418
61	0.8168633/0.098034866 = 8.332376

TABLE 9. SNR_1 FOR THE OPTIMIZED BINARY PHASE–ONLY FILTERS MADE FROM THE OPTIMIZED PHASE–ONLY FILTERS LISTED IN TABLE 7

Number of Tanks	SNR_1 Optimized Binary Phase–Only Filter
1	4.675542/0.086300857 = 54.17723
61	0.6682624/0.1011532 = 6.606438

TABLE 10. SNR_1 FOR THE OPTIMIZED 16-STATE PHASE-ONLY FILTERS MADE FROM THE OPTIMIZED PHASE-ONLY FILTERS LISTED IN TABLE 6

Number of Tanks	SNR_1 Optimized 16-State Phase-Only Filter
1	11.11173/0.053745653 = 206.7466
61	1.515951/0.053154632 = 28.51963

TABLE 11. SNR_1 FOR THE OPTIMIZED 16-STATE PHASE-ONLY FILTERS MADE FROM THE OPTIMIZED PHASE-ONLY FILTERS LISTED IN TABLE 7

Number of Tanks	SNR_1 Optimized 16-State Phase-Only Filter
1	9.839813/0.0098521933 = 998.7434
61	1.139705/0.010111701 = 112.7115

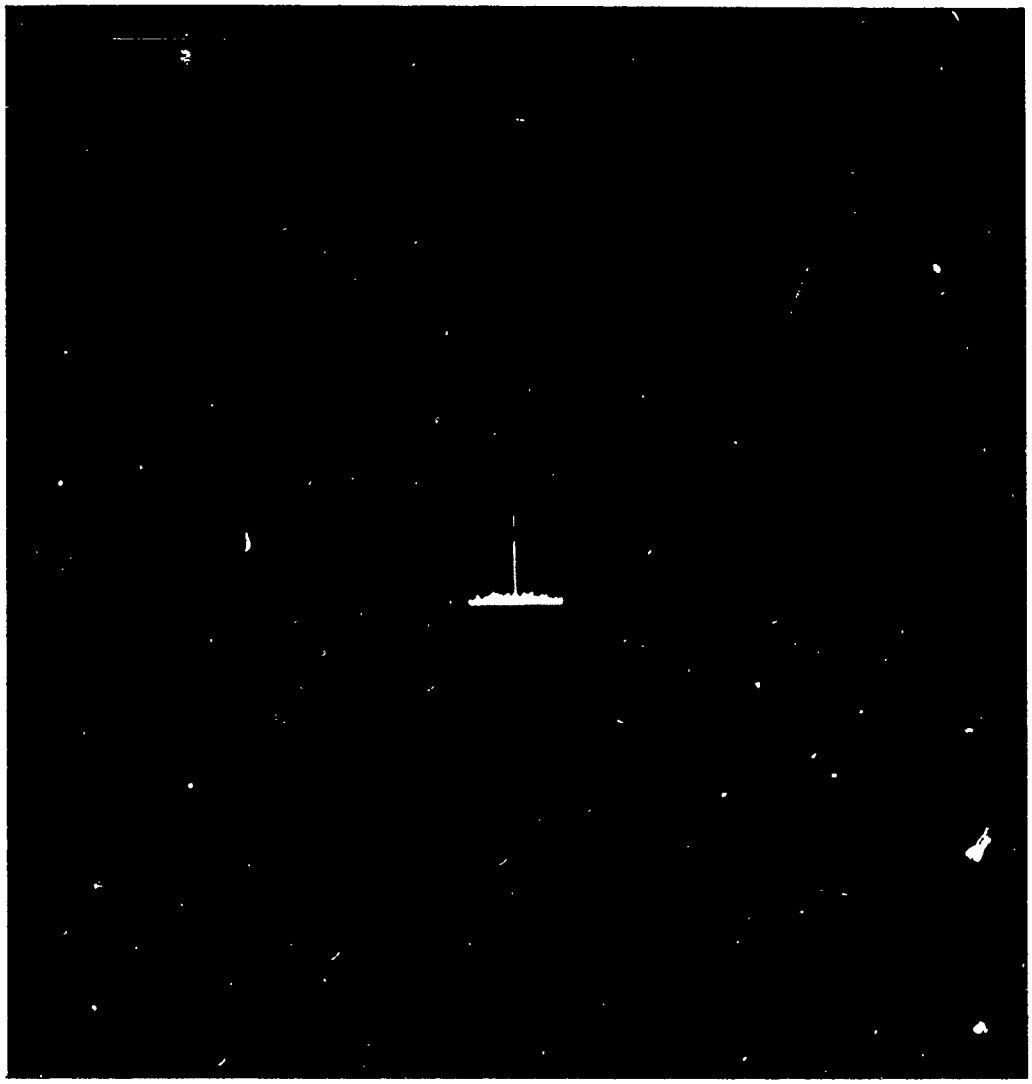


Figure 16. t_4 vs. $\mathcal{F}(h_5)^*$, South-to-North Correlation Plane View

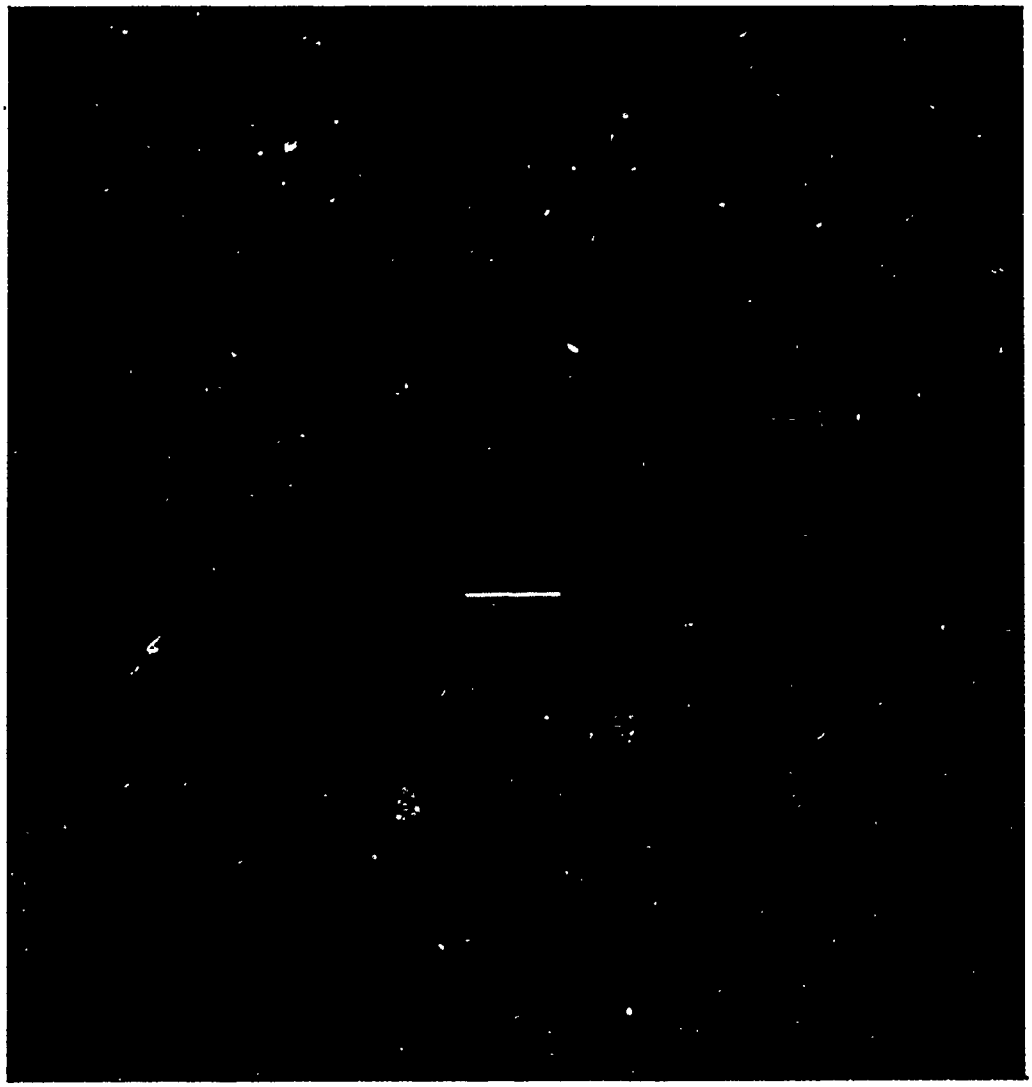


Figure 17. t_5 vs. $\mathcal{F}(h_5)^*$, South–to–North Correlation Plane View

vaguely similar to the T62 training set images. To make an initial test of this, code the M1 image t_6 described in Section VI into a phase-only image, repeat the previous experiments with t_6 in place of t_4, and display the results in Figure 18. There is quite adequate discrimination. A worrisome aspect of this coding of intensity images as phase-only images is that additive noise on the original intensity image is turned into multiplicative noise. How does this new filtering process behave with respect to additive noise on the original amplitude image? To check this, repeat the simulations just discussed with the image t_7 described in Section VI in place of t_4 and display the results in Figure 19. This figure shows that even a great deal of multiplicative noise did not prevent the phase-only filter $\mathcal{F}(h_5)^*$ from recognizing the T62. This is quite gratifying and a bit of a surprise.

Coding intensity images as phase-only images for optical correlator input also seems to work well with small infrared targets of strategic interest. Figure 20 shows a blowup of such a target. Its approximate length is 16 pixels and is presented in a 32 × 32 blank background. Let t_8 be the 32 × 32 image consisting of the target of strategic interest shown in Figure 20 in a uniform background whose intensity is equal to the average intensity on the target. Let $\mathcal{F}(h_6)^*$ be the continuous phase-only filter optimally designed for the training set consisting simply of t_8 as the one true target and a nonzero uniform background as the one false target, optimized over the entire correlation plane. The simulations discussed above were repeated with the image t_8 coded into a phase-only input and with $\mathcal{F}(h_6)^*$ in the filter plane. The resulting south-to-north view of the entire correlation plane is displayed in Figure 21. The filter certainly recognizes the target in spite of the background. This experiment was then repeated with the constant background as input. The ratio of the threshold (the height of the peak signal

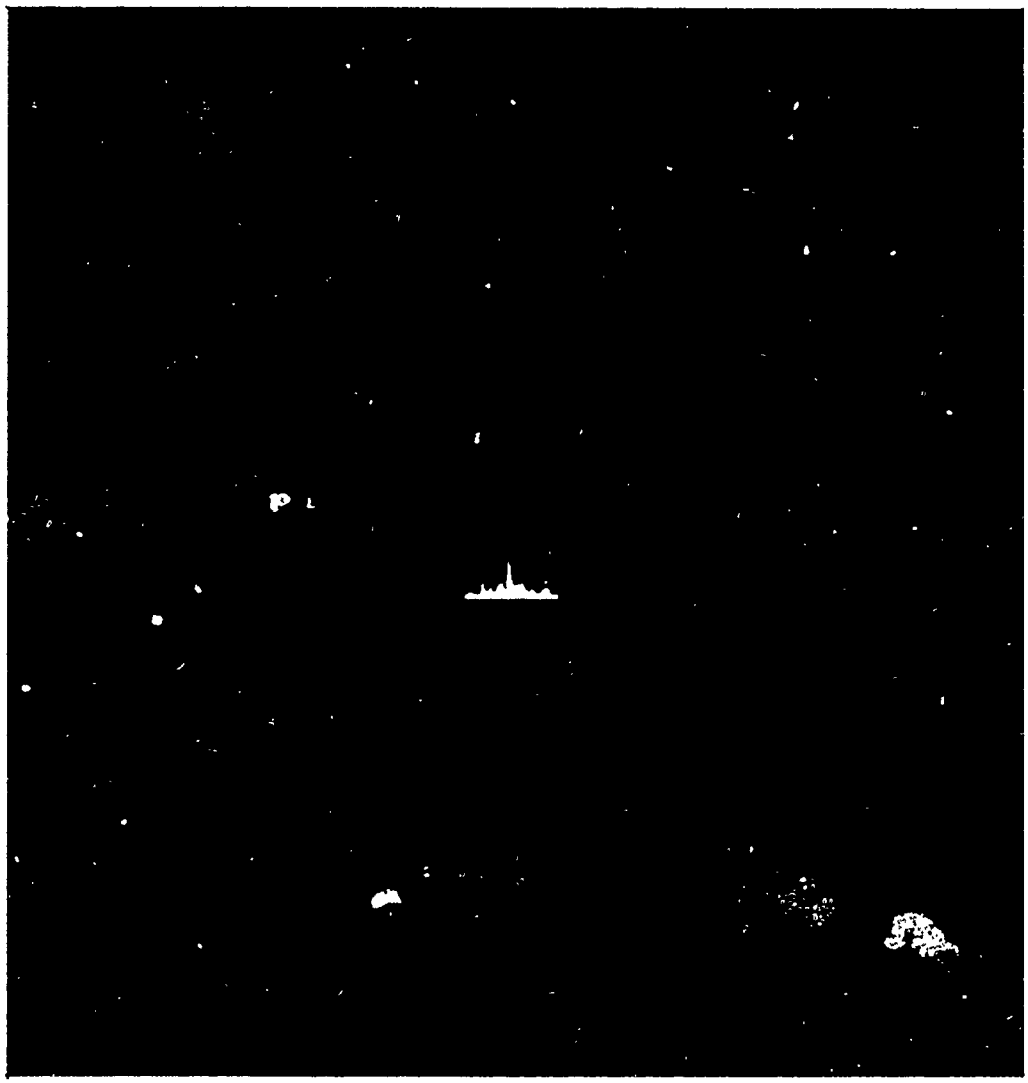


Figure 18. t_6 vs. $\mathcal{F}(h_5)^*$, South-to-North Correlation Plane View

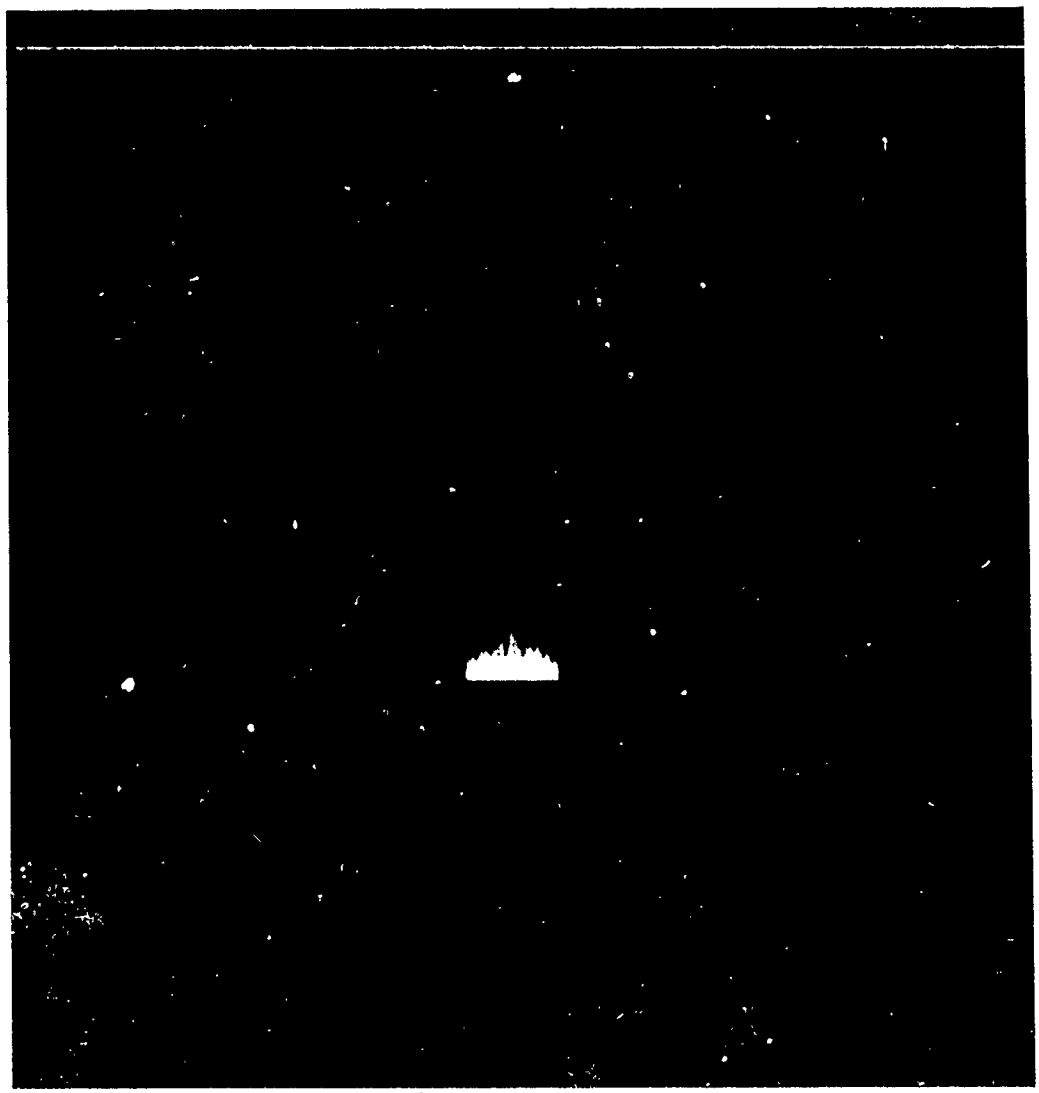


Figure 19. t_7 vs. $\mathcal{F}(h_5)^*$, South-to-North Correlation Plane View

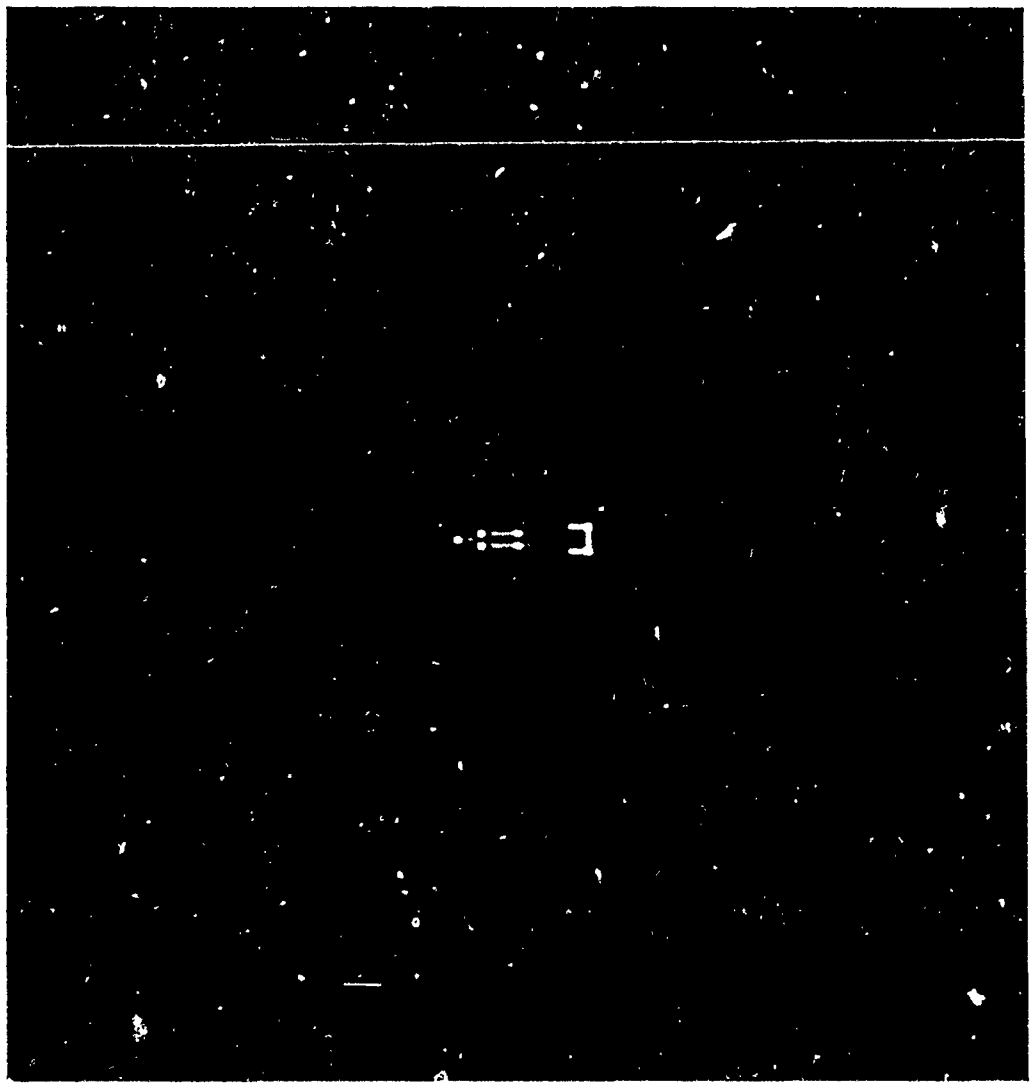


Figure 20. An Infrared Image of a Strategic Target in a Blank Background

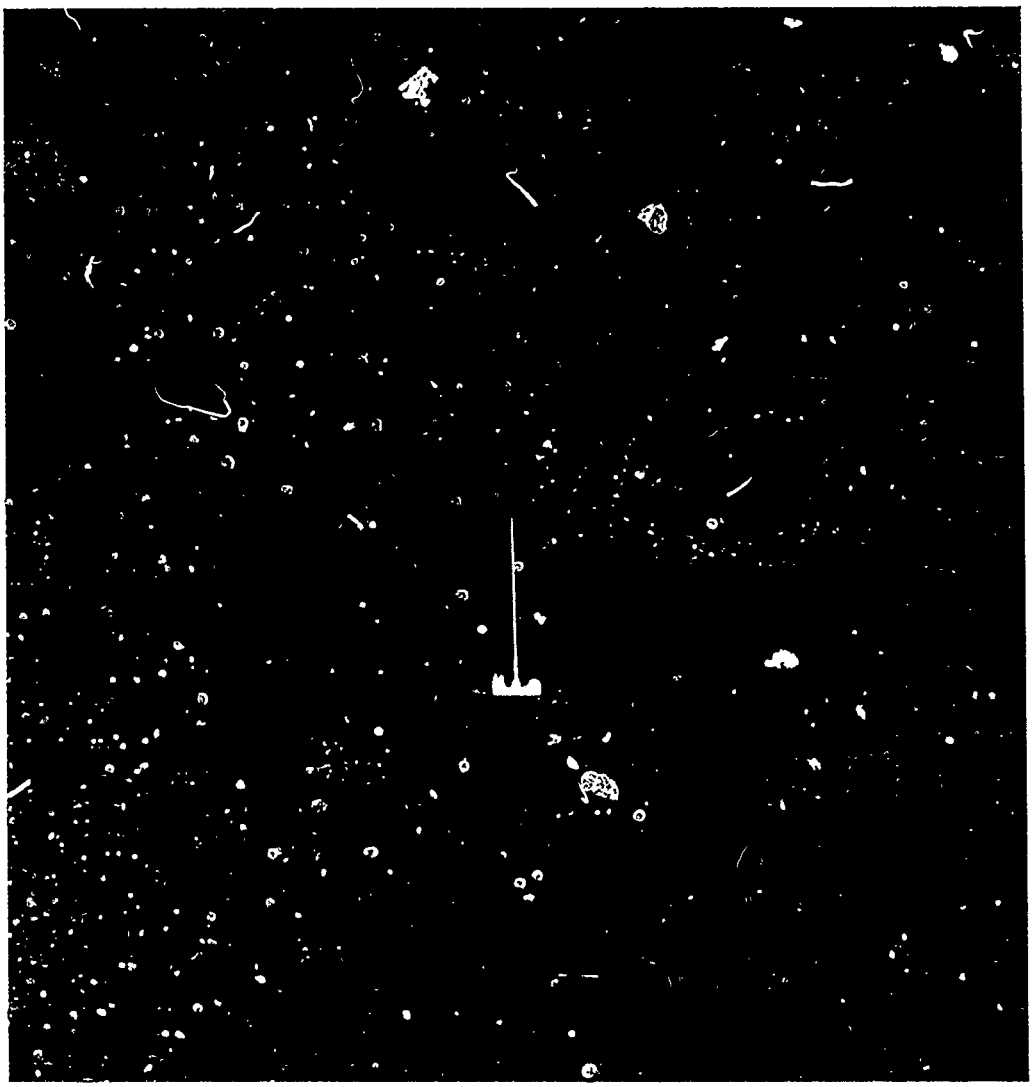


Figure 21. t_8 vs. $\mathcal{F}(h_6)^*$, South-to-North Correlation Plane View

in Figure 21) to the peak signal over the entire correlation plane was 91.43726. This simple example shows the power of the iterative design techniques advocated here, for the corresponding ratio for the phase-only matched filter made from the phase-coded version of t_8 is 0.048271649. This phase-only matched filter was the starting point for the iterative design code for $\mathcal{F}(h_6)^*$. If the filter had been optimized over the detector instead of over the entire correlation plane and this last experiment repeated, then the ratio of the threshold to the peak signal over the detector in the correlation plane would be 1072.659.

SECTION VIII

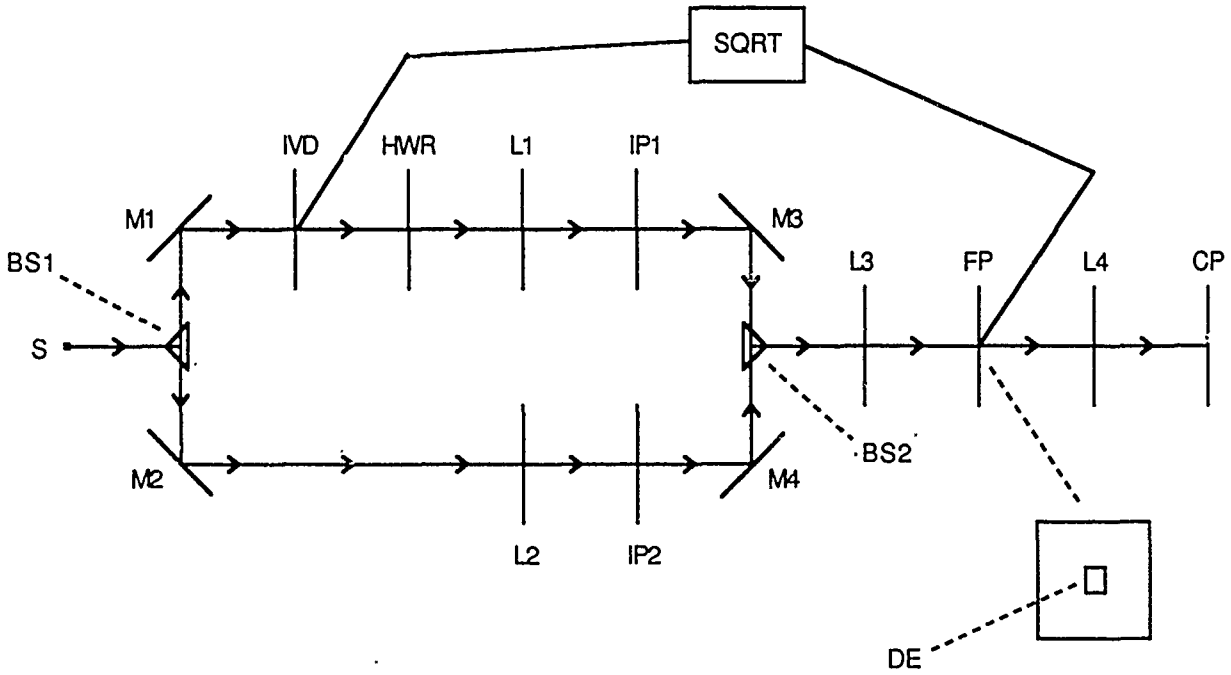
THE THIRD SOLUTION: AN ARCHITECTURE MODIFICATION PLUS PROPER FILTER DESIGN

In this section a perturbation of the standard optical correlator algorithm is discussed which solves Problem #2 of Section V. This algorithm can be implemented in an all-optical system, though at a cost of making the optics more cumbersome than the standard architecture. The ideas of this section are quite intuitive bearing in mind the discussion at the end of Section V. The ideas discussed here are a slight refinement of those previously discussed in Reference 20.

The following is the sequence of operations or algorithm to be performed in carrying out what will be called the all-optical solution to Problem #2 of Section V:

(i) input the coherent image f and measure $|E(f)|$;
(ii) take the Fourier transform $\mathcal{F}(f)$;
(iii) input the coherent image $E(f)\chi_A$ and take the Fourier transform $\mathcal{F}(E(f)\chi_A)$;
(iv) perform the image subtraction $\mathcal{F}(f) - \mathcal{F}(|E(f)|\chi_A) = \mathcal{F}(f - |E(f)|\chi_A)$;
(v) use the spatial light modulator to multiply $\mathcal{F}(f - |E(f)|\chi_A)$ by the filter $\mathcal{F}(h)^*$;
(vi) take the inverse Fourier transform of this product to produce the intensity $|<(f - |E(f)|\chi_A), h_x>|^2$ at each point x of the correlation plane.

$|E(f)|$ is used here rather than $E(f)$ since $E(f)$ need not in general be positive, but in fact may be complex valued (Reference 32). Figure 22 is a schematic of a possible equipment implementation of this algorithm. The next section discusses an



BS1, BS2 = Beam Splitter
CP = Correlation (Output) Plane
DE = Detector Element
FP = Fourier (Filter) Plane
HWR = Half Wave Retarder
IP1 = Input Plane with Square Aperture
IP2 = Image Input Plane
IVD = Intensity Varying Device
L1, L2 = Collimating Lens
L3, L4 = Fourier Transforming Lens
M1, M2, M3, M4 = Mirror
S = Source

Figure 22. A Schematic for a Possible Equipment Implementation of the All-Optical Algorithm

efficient way to measure $|E(f)|$ with a minor perturbation to the correlator architecture. There are some possible drawbacks to this algorithm. For example, it is difficult to conceive of an equipment implementation for this algorithm which does not use two input devices (though the currently fashionable joint transform correlator apparently suffers from the same defect). More importantly, a delicate optical subtraction needs to be performed.

This algorithm for the solution of Problem #2 of Section V requires a new measure of performance for the discriminant function h. In this new context define the signal–to–noise ratio $SNR_2(h)$ of h to be

$$SNR_2(h) = T_2(h)/N_2(h) \tag{4}$$

where $T_2(h)$, the threshold of h, is defined to be

$$T_2(h) = \min_{1 \le i \le n} \max_{x \in B_i} |\langle (f_i - |E(f_i)|\chi_A), h_x \rangle|^2 \tag{5}$$

and where $N_2(h)$, the noise of h, is defined to be

$$N_2(h) = \max_{1 \le i \le m} \max_{x \in R_i - B_i} |\langle (f_i - |E(f_i)|\chi_A), h_x \rangle|^2. \tag{6}$$

In these formulas the f_i's, R_i's, and the B_i's are as described in Section II.

One designs phase–only filters for this all–optical algorithm by numerical techniques similar to those discussed in Section V and The Appendix. If $n = m = 1$, then these numerical techniques essentially produce the phase–only matched filter of Horner and Gianino (Reference 21) applied to the image $f_i - E(f_i)\chi_A$. However, if $m > 1$, this iterative technique differs radically from the technique of Horner and Gianino

(Reference 24) and is necessary for optimal filter design. One can make good binary phase-only filters for the all-optical algorithm from these phase-only filters by a simple perturbation of the one-dimensional optimization techniques discussed in Section II and in Reference 8. Let h_7 be the discriminant function such that $\mathcal{F}(h_7)^*$ is the optimized phase-only filter made from the image t_2 (first discussed in Section IV) for this all-optical algorithm, with its 11 × 11 lowest frequency pixels set equal to 0 as discussed in Section III. Figure 23 shows a south-to-north view of the entire correlation plane of t_2 versus $\mathcal{F}(h_7)^*$. How does this all-optical algorithm perform against targets which are nontrivial and vaguely similar to t_2? Figure 24 shows an M113 from the first series of targets. Let t_9 be the M113 of Figure 24 embedded into a background which is constant and equal to the constant background of t_2. Figure 25 shows a south-to-north view of the entire correlation plane of t_9 versus $\mathcal{F}(h_7)^*$. The horizontal bar in both figures is at the height $T_2(h_7)$. In this instance the all-optical algorithm using the optimized phase-only filter $\mathcal{F}(h_7)^*$ has solved the aperture problem, recognizes the M48, and adequately discriminates against the M113, for in Figure 25, where the entire correlation plane may be regarded as noise, the threshold to peak noise level is 72.0. This is only a simple instance of the efficacy of the all-optical algorithm, but in general the phase-only filters and their optimized binarizations for the all-optical algorithm made by the numerical techniques mentioned here can be packed with a great deal of information, are stable under perturbations in the training set, have a very low false alarm rate, and are extremely insensitive to backgrounds consisting of high variance, high mean, Gaussian or uniform noise. Furthermore, empirical evidence suggests that for the sort of targets used here in a background of zero amplitude, this all-optical algorithm improves

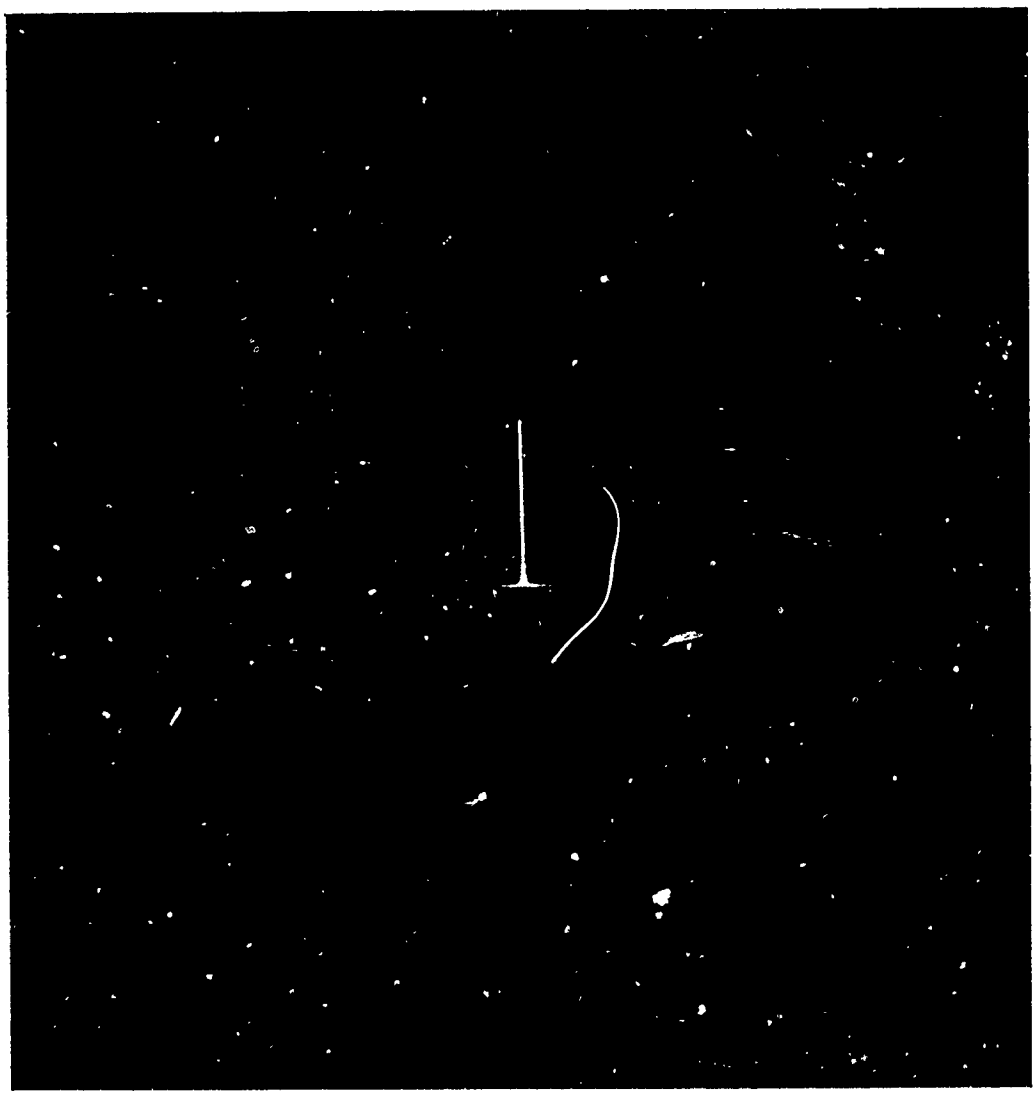


Figure 23. t_2 vs. $\mathcal{F}(h_7)^*$, South–to–North Correlation Plane View

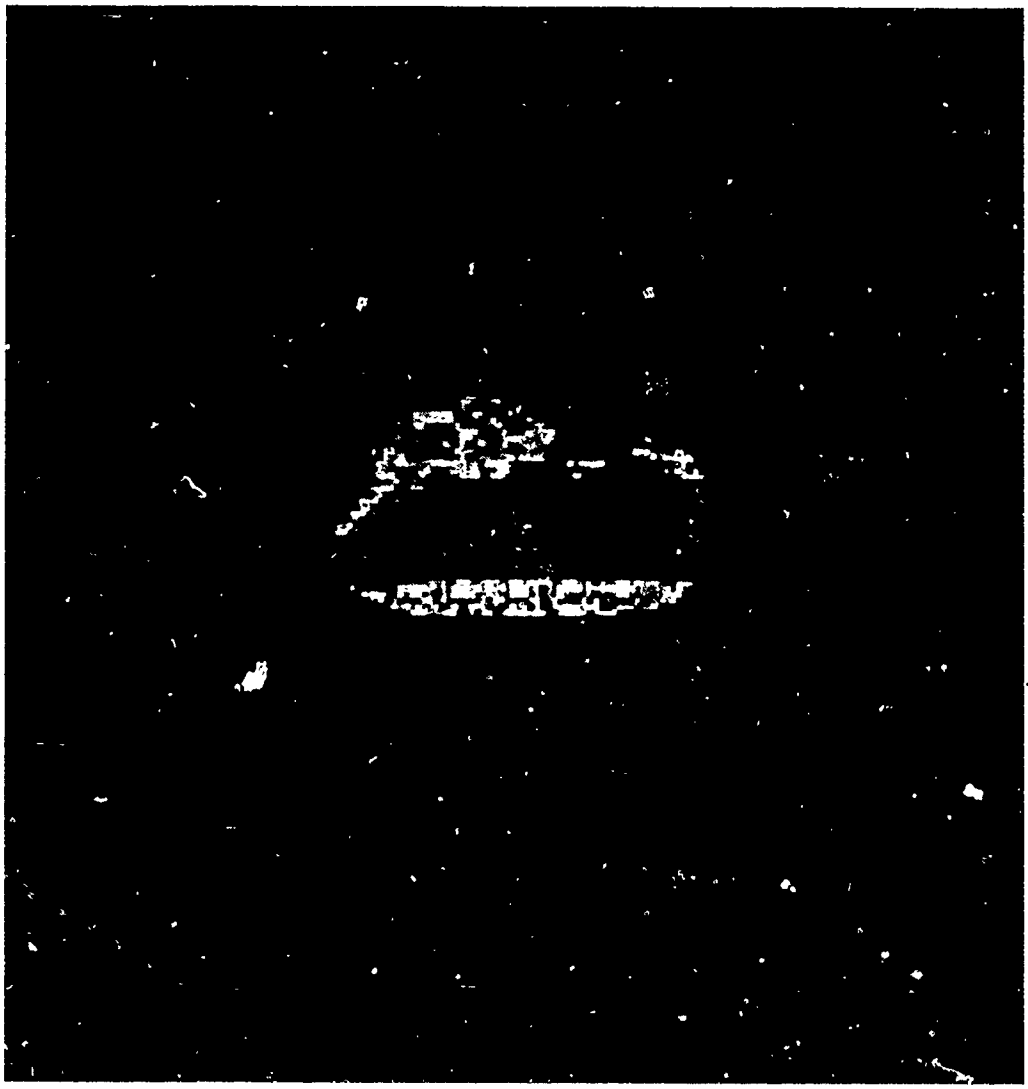


Figure 24. M113 Amplitude Image in Blank Background

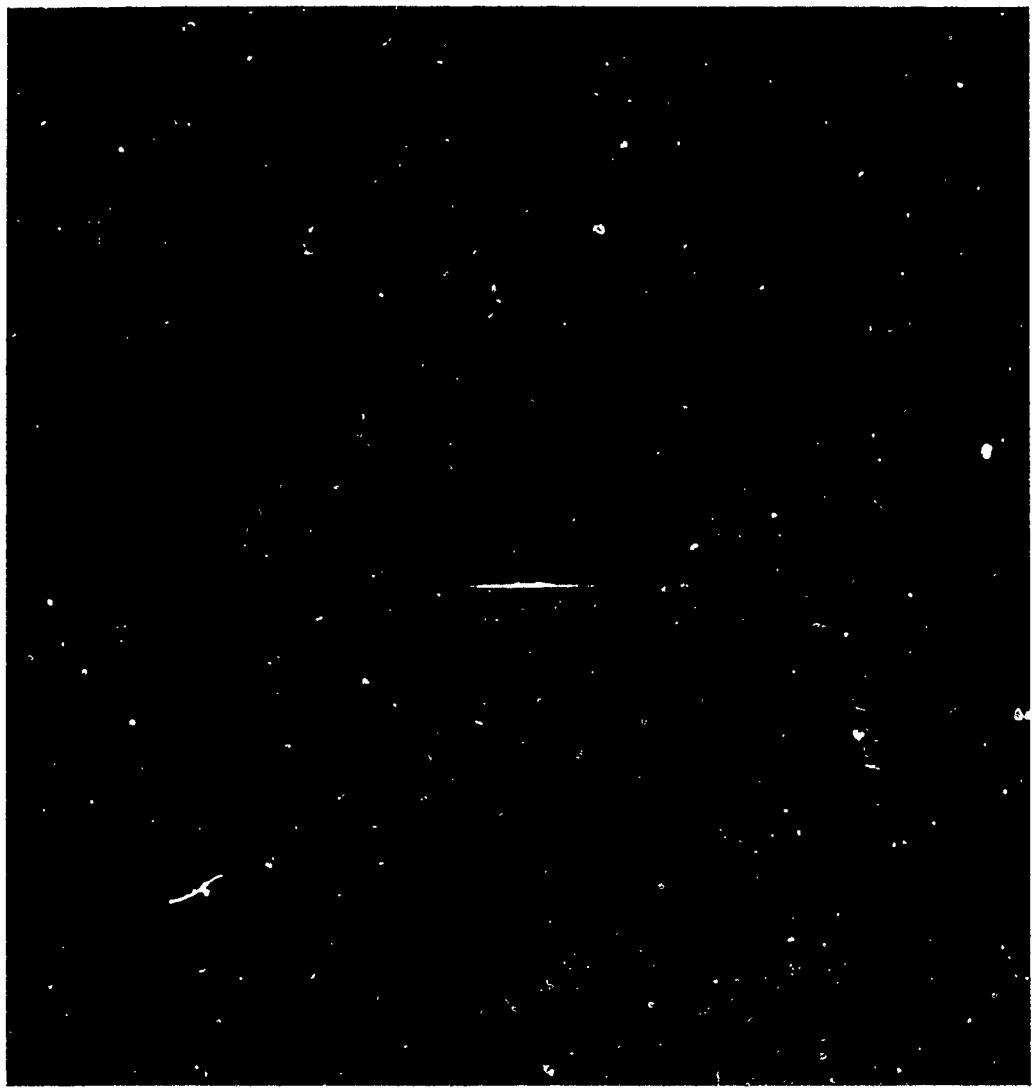


Figure 25. t_9 vs. $\mathcal{F}(h_7)^*$, South-to-North Correlation Plane View

discrimination by more than 25 percent over the standard correlator algorithm.

SECTION IX

THE FOURTH SOLUTION: AN ALGORITHM MODIFICATION PLUS PROPER FILTER DESIGN

This section is devoted to an optical–digital solution of Problem #2 of Section V. It has no advantage over the solutions presented in Section VI and Section VII, but it does have the advantage over the all–optical solution presented in Section VIII in that a major perturbation of the standard optical correlator architecture is avoided, no optical subtractions need be done, and the digital computations needed instead are extremely fast. It unfortunately has the serious flaw that it is not translation invariant. However, it was thought worthwhile to present these ideas, for some correlators (such as the tandem correlator, Reference 34) which lack translation invariance appear to be of interest to the optics community. Furthermore, the translation invariance of this correlator appears to be so unstable that it might well find a possible application in pointing and tracking. A preliminary presentation of the ideas presented in this section was made in Reference 20.

The all–optical algorithm presented in Section VIII may be thought of as a process by which the number

$$|\langle (f - |E(f)|\chi_A), h_x \rangle|^2$$

is stacked over each point x of the input plane. For a carefully crafted h the quantity

$$|\langle (f - |E(f)|\chi_A), h_x \rangle|^2$$

should be relatively large if there is something of interest centered over x and should be relatively small if there is nothing of interest centered near x. But note that

$$|\langle (f - |E(f)|\chi_A), h_x \rangle|^2 \geq (|\langle f, h_x \rangle| - |E(f)| \cdot |\langle \chi_A, h_x \rangle|)^2.$$

The gist of the optical–digital algorithm is to replace the quantities

$$|\langle(f - |E(f)|\chi_A), h_x\rangle|^2$$

with the quantities

$$(|\langle f, h_x\rangle| - |E(f)|\cdot|\langle\chi_A, h_x\rangle|)^2$$

in this stacking process. There is no apparent physical interpretation for this idea, but it was suggested by purely mathematical considerations, the success of the all–optical algorithm presented in Section VIII, the desire to eliminate the need for optical subtractions, the desire to reserve for optics what optics does best, viz., correlations, and the desire to reserve for digital calculations only very simple fast operations.

The following is the sequence of operations to be in carrying out the optical–digital algorithm:

(i) input the image χ_A and measure and store the quantities $|\langle\chi_A, h_x\rangle|^2$;

(ii) input the image f, take its Fourier transform $\mathcal{F}(f)$, measure $|E(f)|^2 = C|\mathcal{F}(f)(0,0)|^2$, store $|E(f)|^2$, and compute and store $|E(f)|$;

(iii) measure and store the quantities $|\langle f, h_x\rangle|^2$;

(iv) digitally compute the quantities

$$|\langle f, h_x\rangle|^2 + |E(f)|^2\cdot|\langle\chi_A, h_x\rangle|^2 - 2|E(f)|\cdot(|\langle f, h_x\rangle|^2\cdot|\langle\chi_A, h_x\rangle|^2)^{(1/2)}$$

$$= (|\langle f, h_x\rangle| - |E(f)|\cdot|\langle\chi_A, h_x\rangle|)^2.$$

This optical–digital algorithm can be implemented in the standard correlator architecture, supplemented by some digital electronics, with only one minor change. Step (ii) requires the

measurement of the quantity $|\mathcal{F}(f)(0,0)|^2$. This can be done by placing a detector element in the middle of the spatial light modulator. As was pointed out before, some of the low frequency pixels in these phase-only filters were set equal to 0. Therefore, this detector element will not interfere with the implementation of the these phase-only filters. Up to a universal constant depending on the design of the correlator but independent of f, $\mathcal{F}(f)(0,0)$ is a scalar multiple of E(f). So there indeed is a positive universal constant C, dependent on the design of the correlator but independent of f, so that $|E(f)|^2 = C|\mathcal{F}(f)(0,0)|^2$. C can be theoretically calculated and its value checked by measurements. The digital computations required in step (iv) should proceed at a rapid pace. Step (iv) was written in what appears to be a rather awkward sequence of substeps, but the latter are designed to make maximal use of the measured stored quantities and to minimize the number of square root operations. The square root operations might be a potential bottleneck, but it is plausible that state-of-the art specialized square roots chips will allow this algorithm to proceed at a frame rate of several thousand per second for a 128 × 128 device. Furthermore, for purposes of recognition and discrimination, it is probable that one will need to compute the square roots in step (iv) with only two significant places of accuracy. This can be accomplished by simple lookup table procedures and should allow the digital portion of the optical-digital algorithm to proceed at whatever pace is needed. Use the term optical-digital architecture to denote the standard optical correlator architecture with the very slight modification discussed here.

The optical-digital algorithm requires a new measure of performance of the discriminant function h. In this new context define the signal-to-noise ratio of h to be

$$SNR_3(h) = T_3(h)/N_3(h) \tag{7}$$

where $T_3(h)$, the threshold of h, is defined to be

$$T_3(h) = \min_{1 \le i \le n} \max_{x \in B_i} (|\langle f_i, h_x \rangle| - |E(f_i)| \cdot |\langle \chi_A, h_x \rangle|)^2 \qquad (8)$$

and where $N_3(h)$, the noise of h, is defined to be

$$N_3(h) = \max_{1 \le i \le m} \max_{x \in R_i - B_i} (|\langle f_i, h_x \rangle| - |E(f_i)| \cdot |\langle \chi_A, h_x \rangle|)^2. \qquad (9)$$

In these formulas the f_i's, R_i's, and the B_i's are as described in Section II.

$SNR_3(h)$ and $T_3(h)$ are even more nondifferentiable than are $SNR_1(h)$, $SNR_2(h)$, $T_1(h)$, and $T_2(h)$. However, the mathematical techniques and algorithms described in Section V and The Appendix work in this case also. One designs phase-only filters for this optical-digital algorithm by methods that are philosophically similar to but technically more difficult than the methods discussed in previous sections. One can make good binary phase-only filters for the optical-digital algorithm from these phase-only filters by a simple perturbation of the one-dimensional optimization techniques discussed in Section II and in Reference 8. Let h_8 be the discriminant function such that $\mathcal{F}(h_8)^*$ is the optimized phase-only filter made from t_2 for the optical-digital algorithm by the methods discussed here. Figure 26 shows a south-to-north view of the entire correlation plane of t_2 versus $\mathcal{F}(h_8)^*$, and Figure 27 shows a south-to-north view of the entire correlation plane of t_9 versus $\mathcal{F}(h_8)^*$. The horizontal bar in both figures is at the height $T_3(h_8)$. In this instance the optical-digital algorithm has solved the aperture problem, recognizes the M48, and adequately discriminates against

Figure 26. t_2 vs. $\mathcal{F}(h_8)^*$, South-to-North Correlation Plane View

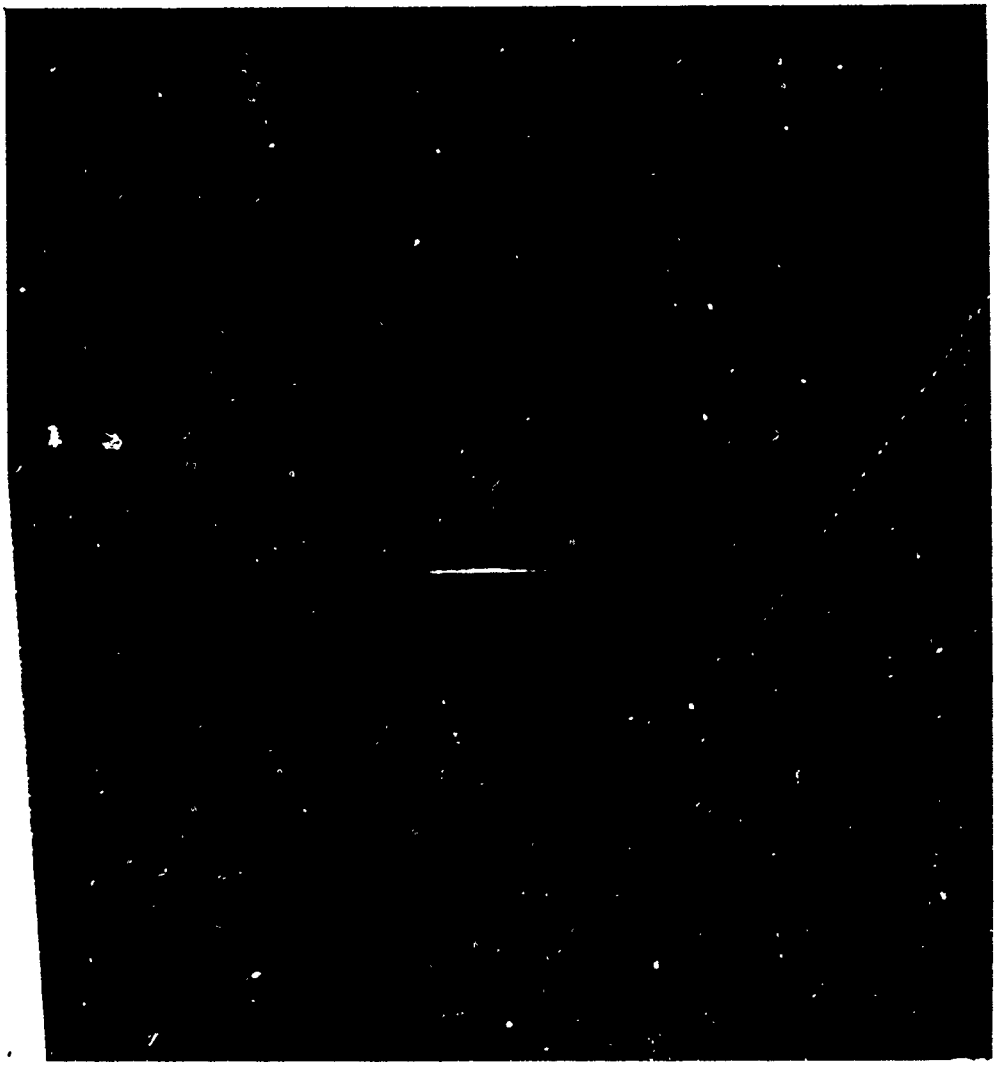


Figure 27. t_9 vs. $\mathcal{F}(h_8)^*$, South-to-North Correlation Plane View

TABLE 12. STABILITY OF $\mathcal{F}(h_9)^*$ VERSUS THE M48 IMAGES

M48 Image	Peak Signal/$T(h_9)$
1	1.0000150
2	1.0000675
3	1.0000921
4	1.0000422
5	1.0001456
6	1.0000572
7	1.0053832
8	1.0000864
9	1.0000509
10	1.0000657
11	1.0001077
12	1.0000860
13	1.0000537
14	1.0000387
15	1.0155759
16	1.0000606
17	1.0000454
18	1.0378693
19	1.0000432
20	1.0000947
21	1.0000000

no information about the M113 images was used in the construction of h_9. Carefully designed binary phase-only filters work well in this optical-digital architecture. Use an obvious perturbation of the one-dimensional optimization techniques discussed in Section II applied to h_9 to obtain h_{10}, a discriminant function such that $\mathcal{F}(h_{10})^*$ is a binary phase-only filter. This optimization process is important here, for it leads to a 25 percent increase in $T_3(h_{10})$ and a more than doubling of $SNR_3(h_{10})$ over arbitrary techniques. Table 13 illustrates just how stable a peak signal may be obtained if one uses the numerical techniques described here to construct h_{10}. Furthermore, the ratio of $T_3(h_{10})$ divided by the maximum peak signal off any of the 21 M113's is 24.2. The average over the 21 M113 images of $T_3(h_{10})$ divided by the peak signal off each M113 is 33.4. Thus h_{10} discriminates the M48 images quite well from the M113 images. Note again that there was this strong signal stability and good discrimination even though no information about the M113 images was used in the construction of h_{10}. Further simulations show that h_9 and h_{10} are extremely insensitive to backgrounds consisting of high variance, high mean, Gaussian or uniform noise. Empirical evidence suggests that for the sort of targets used here in a background of zero amplitude, this optical-digital algorithm improves discrimination by more than 25 percent over the standard correlator algorithm. As a test of the power and robustness of the numerical techniques described here to make filters for the optical-digital algorithm, a number of tests were run starting with a training set consisting of the 21 M48 images from the first series in a constant background whose amplitude was the average of nonzero pixels on the M48's themselves and using a randomly generated 128 × 128 phase-only matrix as the initial point in the iterative process. When completed, the algorithms described in this report had increased the threshold by a factor of 1.6×10^8 to a value within a few percent of that

TABLE 13. STABILITY OF $\mathcal{F}(h_{10})^*$ VERSUS THE M48 IMAGES

M48 Image	Peak Signal/$T(h_{10})$
1	1.026976
2	1.014516
3	1.010611
4	1.019399
5	1.011869
6	1.016052
7	1.011435
8	1.014895
9	1.012778
10	1.026629
11	1.033311
12	1.020809
13	1.017628
14	1.011248
15	1.024196
16	1.000000
17	1.008831
18	1.030585
19	1.012348
20	1.011092
21	1.013160

of h_9. The resulting phase-only filter for the optical-digital architecture gave qualitatively the same signal stability and discrimination characteristics as $\mathcal{F}(h_9)^*$.

The algorithm discussed in this section is just one of a whole family of possible filtering algorithms which can be used in connection with the optical-digital architecture. In fact, let $\varphi(u,v)$ be any nonnegative function defined on the first quadrant ($u \geq 0$, $v \geq 0$) of (u,v)-space with the property that $\varphi(u,u) = 0$ if $u \geq 0$. Use the optical-digital architecture to compute and store the quantities $|E(f)|^2$, $|\langle f,h_x\rangle|^2$, and $|\langle \chi_A,h_x\rangle|^2$. Over each point x of the input plane stack the number $\varphi(|\langle f,h_x\rangle|^2, |E(f)|^2 \cdot |\langle \chi_A,h_x\rangle|^2)$. If φ permits, one then needs to carefully design a discriminant function h so that $\mathcal{F}(h)^*$ is a phase-only filter, so that $\varphi(|\langle f,h_x\rangle|^2, |E(f)|^2 \cdot |\langle \chi_A,h_x\rangle|^2)$ is relatively large if there is a target of interest centered at x, and so that $\varphi(|\langle f,h_x\rangle|^2, |E(f)|^2 \cdot |\langle \chi_A,h_x\rangle|^2)$ is relatively small if there is no target centered close to x. There are obvious analogs of Equations (7) - (9) to measure the performance of h for this optical-digital architecture and φ-dependent filtering algorithm. To optimally design h, one would have to devise φ-dependent iterative algorithms similar to those described in here. $\varphi(u,v) = (u^{1/2} - v^{1/2})^2$ for the case considered above. Another choice for φ which immediately springs to mind is $\varphi(u,v) = |u^2 - v^2|$. This φ is attractive, for the digital processing needed to compute it needs no square root operations. The program sketched in this paragraph has been carried out for $|u^2 - v^2|$, but it was found to provide only half of the discrimination capability supplied by $(u^{1/2} - v^{1/2})^2$. This intuitively may be traced to the observation that an extraneous cross term is approximately canceled if one chooses φ

to be $(u^{1/2} - v^{1/2})^2$ rather than $|u^2 - v^2|$.

SECTION X

CONCLUSIONS

There were two objectives to this program. First, develop mathematical techniques and algorithms to solve the complicated, nonlinear, nondifferentiable minimax questions in a large number of variables which arise in the formulation of phase-only filter design given in Section II. Implement these techniques and algorithms into effective and efficient computer codes. Second, use these mathematical techniques, algorithms, and computer codes to resolve the background problem in target recognition described in detail in Section IV. These objectives have in large measure been accomplished. The mathematical techniques are described in Section II, Section V, and Appendix A, and their implementation into computer codes is discussed in Section V. Four solutions to the background problem were given in Section VI, VII, VIII, and IX. The solution discussed in Section VI is the current favorite. The approach and techniques discussed in Section VI should be tested in simulations over a wide range of training sets and the resulting filters implemented and tested in actual equipment. The approach and results presented in Section VII are very new and thus very preliminary. Nonetheless, the results of this new approach are so encouraging that a large effort is justified to test it exhaustively. It is questionable if the solution presented in Section VIII is practical for applications in the military sphere, for it is difficult to envision an equipment implementation which does not employ two input devices (but, on the other hand, the currently fashionable joint transform correlator also needs two input devices) and, in addition, a delicate optical subtraction needs to be performed. The methods presented in Section VIII might find some commercial and/or manufacturing use, but an effort should be expended in its

study only if the ideas advocated in Section VI and Section VII fail. The solution presented in Section IX is not useful for target recognition since it is not translation invariant, but it might prove useful in a pointing and tracking application.

THE APPENDIX

THE CHOICE OF AN OPTIMAL SEARCH DIRECTION AS A GEOMETRIC PROBLEM

The contents of this mathematical appendix certainly must be well known, but it is difficult to give a reference for a coherent discussion of the topics considered here. Optimization questions of various sorts are familiar to anyone who has taken a calculus course or an operations research course. The sort of optimization question of interest in this study does not seem to fall neatly into either of these subject areas, but rather involves aspects of both. All of the optimization problems discussed in previous sections can be cast into the following general form. Let X be a nonempty open subset of the k-dimensional Euclidean space R^k. Let $\varphi_1, \ldots, \varphi_N$ be a finite sequence of differentiable functions on X which are bounded below. Define a function φ on X by setting

$$\varphi(x) = \min_{1 \leq i \leq N} \varphi_i(x)$$

for each x in X. Problem: Find an iterative algorithm to drive φ to a local maximum on X. This problem is somewhat involved, for even though each φ_i may be as differentiable as could be desired, φ itself almost always is not. This is not merely a pedantic point, for all of the difficulty in optimizing φ occurs at those points where φ is not differentiable. Hence, simple steepest ascent methods do not apply here. In the problems considered in previous sections, φ is actually a function on a very large torus or product of circles and not an open subset of some Euclidean space. However, any such torus is locally parameterized by an open subset of some Euclidean space, and the discussion here encompasses the problems considered in previous sections. In the previous sections a typical size for k is 128^2 –

$11^2 = 16{,}263$ and N can vary between the relatively small number of images in the training set when optimizing a threshold to a number in the millions when optimizing a signal-to-noise ratio. Because of the large number of functions involved and the large number of variables, it is important to use algorithms in which no large sets of linear equations need be solved or large matrices inverted. The algorithm sketched below has this property.

Suppose a point x_0 in X is given. An iterative scheme involves the choice of a feasible direction (or unit vector) d in R^k so that $\varphi(x_0) < \varphi(x_0 + \alpha d)$ for some sufficiently small positive number α. Given that such a direction d exists, it is desired to choose the best possible d. If no such d exists, then the iteration stops. In order to determine whether or not such a d exists, one proceeds as follows. Let F be the set of all integers i between 1 and N such that i is active—such that $\varphi_i(x_0)$ is very close to $\varphi(x_0)$. Sometimes, particularly at the beginning of an iteration, the size of F might be 1, but in practice it quickly grows to be rather large. The size of F of course depends on what tolerance is used to determine when $\varphi_i(x_0)$ is very close to $\varphi(x_0)$. This tolerance usually should go to 0 as the iteration proceeds. Calculate the gradients $v_j = (\nabla\varphi_j)(x_0)$ $(j \in F)$. A direction d is feasible provided each dot product $v_j \cdot d$ is positive. The best direction d to choose is the solution to

$$\max_{\substack{d \in R^k \\ |d| = 1}} \min_{j \in F} v_j \cdot d \qquad \text{(A-1)}$$

where $|x|$ is the length of the vector $x \in R^k$. Apparently one intractable problem has been replaced with another. However, this is not the case, for the vector d which solves Equation

(A–1) admits an interesting geometric interpretation.

Let P be the polytope in R^k spanned by the vectors v_j ($j \in F$). P in fact is the convex hull of the vectors v_j ($j \in F$), viz.,

$$P = [\ \sum_{j \in F} t_j v_j \mid t_j \geq 0,\ \sum_{j \in F} t_j = 1\ ].$$

Note that if the origin $0 \in P$, then there is no unit vector d such that $v_j \cdot d > 0$ ($j \in F$). For if $v_j \cdot d > 0$ ($j \in F$), then $v \cdot d > 0$ for all $v \in P$. But $0 \cdot d = 0$. This is an obvious contradiction, and so there is no unit vector d such that $v_j \cdot d > 0$ ($j \in F$) if $0 \in P$. In this case either the tolerance used in determining F must be shrunk or the algorithm must terminate, for there is no feasible direction in which to move to try and improve φ. So it may be supposed that $0 \notin P$. P is a compact convex subset of R^k and therefore contains a unique point $w \in P$ which is closest to but distinct from 0. Then $d = w/|w|$ is the unique solution to Equation (A–1). To prove this, note that

$$|tv + (1 - t)w|^2 \geq |w|^2$$

for all $0 < t < 1$ since P is convex and w is the point of P with smallest length. Expanding, this implies that

$$t^2 v \cdot v + 2t(1 - t)v \cdot w \geq t(2 - t)w \cdot w.$$

Divide this expression by 2t and then let $t \to 0+$. One obtains that

$$v \cdot w \geq w \cdot w$$

for all $v \in P$. This implies that

$$v \cdot d \geq w \cdot d = |w|$$

for all $v \in P$. This in turn implies that

$$w \cdot d = |w| = \min_{j \in F} v_j \cdot d$$

since w is a convex combination of the v_j ($j \in F$). Now let d' be any other unit vector distinct from $d = w/|w|$. Then

$$\min_{j \in F} v_j \cdot d' \leq w \cdot d' < |w| \cdot |d'| = |w| = \min_{j \in F} v_j \cdot d$$

by the strict form of the Cauchy–Schwarz inequality. Hence,

$d = w/|w|$ is the unique solution to Equation (A–1).

The above analysis implies that one can find the best direction $d = w/|w|$ to move by finding the

$$w = \sum_{j \in F} t_j v_j$$

which solves the quadratic programming problem:

$$\text{minimize } \left| \sum_{j \in F} t_j v_j \right|^2$$

subject to the constraints $t_j \geq 0$, $\sum_{j \in F} t_j = 1$.

Speed is of the essence in solving this problem, for it is only an intermediate problem in each iterative step. If the size of F is two or three, then a good approximate solution can be found by Monte Carlo methods. This is certainly not a cure–all, for heavy handed reasoning shows that if one wants an approximate solution within δ of the true solution and the size of F is p, then one will have to make approximately $(1/\delta)^p$ iterations. This quadratic programming problem can be solved by standard quadratic programming techniques which are given in Dantzig, Reference 35. This technique is relatively slow and in essence involves solving systems of linear equations. Techniques which avoid this are to be preferred because the size of F can be large. There appear to be very fast iterative algorithms for solving this polytope problem (less than a second on a VAX 11–785, even if the number of vectors is 250 or more) which usually give at least an order of magnitude improvement in speed in most cases of interest over any other techniques. Algorithms which incorporate such iterative algorithms have been incorporated into the design codes discussed in this report.

REFERENCES

1. S.H. Lee, "Basic Principles of Optical Information Processing," Optical Information Processing Fundamentals, Springer-Verlag, 1981.
2. H.J. Weaver, Applications of Discrete and Continuous Fourier Analysis, Wiley-Interscience, 1983.
3. J.W. Goodman, Introduction to Fourier Optics, McGraw-Hill, 1968.
4. G.O. Reynolds, J.B. DeVelis, G.B. Parrent, Jr., and B.J. Thompson, The New Physical Optics Notebook: Tutorials in Fourier Optics, copublished by SPIE - The International Society for Optical Engineering and the American Institute of Physics, 1989.
5. M. Born and E. Wolf, Principles of Optics, sixth (corrected) edition, Pergamon Press, 1987.
6. J.D. Gaskill, Linear Systems, Fourier Transforms, and Optics, Wiley, 1978.
7. D. Casasent and H.J. Caulfield, "Basic Concepts," Optical Data Processing Applications, Springer-Verlag, 1978.
8. R.R. Kallman, "Optimal Low Noise Phase-Only and Binary Phase-Only Correlation Filters for Threshold Detectors," Applied Optics, Vol 25, Page 4216, 1986.
9. M.I. Skolnik, Radar Handbook, McGraw-Hill, 1970.
10. D.L. Flannery, A.M. Biernacki, J.S. Loomis, and S.L. Cartwright, "Real-Time Coherent Correlator Using Binary Magnetooptic Spatial Light Modulators at Input and Fourier Planes," Applied Optics, Vol 25, Page 466, 1986.
11. J.M. Florence, M.K. Giles, and J.Z. Smith, "Operation of a Deformable Mirror Device as a Fourier Plane Modulating Filter," Digital and Optical Shape Representation and Pattern Recognition, SPIE, 1988.

12. P. Keller, D.L. Flannery, S.L. Cartwright, and J.S. Loomis, "Performance of Binary Phase-Only Correlation on Machine Vision Imagery," Optics, Illumination, and Image Sensing for Machine Vision, SPIE, 1986.
13. D.R. Pape and L.J. Hornbeck, "Characteristics of the Deformable Mirror Device for Optical Information Processing," Optical Engineering, Vol 22, Page 675, 1983.
14. D. Psaltis, E.G. Paek, and S.S. Venkatesh, "Optical Image Correlation with a Binary Spatial Light Modulator," Optical Engineering, Vol 23, Page 698, 1984.
15. W.E. Ross, D. Psaltis, and R.H. Anderson, "Two-Dimensional Magneto-Optic Spatial Light Modulator for Signal Processing," Optical Engineering, Vol 22, Page 485, 1983.
16. B.V.K. Vijaya Kumar and Z. Bahri, "Phase-Only Filters with Improved Signal to Noise Ratio," Applied Optics, Vol 28, Page 250, 1988.
17. D. Flannery, J. Loomis, and M. Milkovich, "New Formulations for Discrete-Valued Correlation Filters," Digital and Optical Shape Representation and Pattern Recognition, SPIE, 1988.
18. D.L. Flannery, J.S. Loomis, and M.E. Milkovich, "Transform Ratio Ternary Phase-Amplitude Filter Formulation for Improved Correlation Discrimination," Applied Optics, Vol 27, Page 4079, 1988.
19. M.E. Milkovich, D.L. Flannery, and J.S. Loomis, "Transform-Ratio Ternary Phase-Amplitude Filter Formulations for Character Recognition," Optical Engineering, Vol 28, Page 487, 1989.
20. R.R. Kallman, "Two Variants of the Optical Correlation Process," Digital and Optical Shape Representation and Pattern Recognition, SPIE, 1988.

21. J.L. Horner and P.D. Gianino, "Phase-Only Matched Filtering," Applied Optics, Vol 23, Page 812, 1984.
22. R.R. Kallman, "Direct Construction of Phase-Only Correlation Filters," Real Time Signal Processing X, SPIE, 1987.
23. R.R. Kallman, "Direct Construction of Phase-Only Filters," Applied Optics, Vol 26, Page 5200, 1987.
24. J.L. Horner and P.D. Gianino, "Applying the Phase-Only Filter Concept to the Synthetic Discriminant Correlation Filter," Applied Optics, Vol 24, Page 851, 1985.
25. M. Jones, "INMOS Marketing Update," Occam User Group Newsletter, No 12, Page 83, 1990.
26. R.R. Kallman, "The Optimal Construction of Synthetic Discriminant Functions for Optical Matched Filters," U.S. Air Force Summer Faculty Research Program, 1984 Technical Reports, Vol II, SCEEE Press, 1985.
27. R.R. Kallman, The Optimal Construction of Synthetic Discriminant Functions (SDF's), AFATL-TR-86-19, Air Force Armament Laboratory, Eglin Air Force Base, Florida 32542, April 1986.
28. R.R. Kallman, The Construction of Low Noise Optical Correlation Filters and Their Application to Target Identification Problems, AFATL-TR-86-20, Air Force Armament Laboratory, Eglin Air Force, Florida 32542, July 1986.
29. A.V. Oppenheim, "The Importance of Phase in Signals," Proceedings of the IEEE, Vol 69, Page 529, 1981.
30. N. Konforti, E. Marom, and S.-T. Wu, "Phase-Only Modulation with Twisted Nematic Liquid-Crystal Spatial Light Modulators," Optics Letters, Vol 13, Page 251, March 1988.
31. T.H. Barnes, T. Eiju, K. Matsuda, and N. Ooyama, "Phase-Only Modulation Using a Twisted Nematic Liquid Crystal Television," Applied Optics, Vol 28, Page 4845, 1989.

32. A.D. Gara, "Phase Response of a Liquid Crystal Image Transducer," Applied Optics, Vol 17, Page 3696, 1978.
33. H.O. Bartlet, Computer-Generated Holographic Component with Optimum Light Efficiency," Applied Optics, Vol 23, Page 1499, 1984.
34. H.O. Bartlet, "Applications of the Tandem Component: an Element with Optimum Light Efficiency," Applied Optics, Vol 24, Page 3811, 1985.
35. G.B. Dantzig, Linear Programming and Extensions, Princeton University Press, 1963.